IEEE ICAS 2021

INTERNATIONAL CONFERENCE ON AUTONOMOUS SYSTEMS

Virtual Conference | August 11-13,2021

- Welcome Message
- Table of Contents
- Technical Papers
- Authors Index

2021 CONFERENCE PROCEEDINGS

Please visit website for more information!

2021.ieee-icas.org

SPONSORS AND ORGANIZERS

ISBN: 978-1-7281-7289-7
Part Number: CFP21GAF-ART

Technical Support

Conference Catalysts

Phone: +1 352 872 5544

cdyer@conferencecatalysts.com

2021 IEEE International Conference on Autonomous Systems (ICAS) Proceedings

Additional copies may be ordered from:
IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08855-1331 USA

+1 800 678 IEEE (+1 800 678 4333)
+1 732 981 1393
+1 732 981 9667 (FAX)
email: customer-service@ieee.org

IEEE Part Number: CFP21GAF-ART
ISBN: 978-1-7281-7289-7

Welcome Message from the Chairs

Dear Colleagues,

It gives us great pleasure to welcome you all to the inaugural 2021 IEEE International Conference on Autonomous Systems (IEEE ICAS 2021). The IEEE ICAS conference is sponsored by IEEE Signal Processing Society (SPS) through its IEEE SPS Autonomous Systems Initiative (ASI). The Organizing Committee has worked hard to develop a technical program that focuses on advanced Signal Processing and Machine Learning theories and techniques relevant to the development of next generation autonomous systems. Over the past decade, researchers have proposed computing systems with advanced levels of autonomy to manage ever-increasing complex tasks. Signal Processing plays a key role by developing underlying theories and techniques necessary to design algorithms used for perception, control and learning within such autonomous system.

In 2019, we started the journey to organize the inaugural IEEE ICAS event at Concordia University, Montreal, Canada in August 2020. The COVID-19 pandemic, which has changed the world as we know it, forced us to make the hard decision of postponing the conference to 2021. We were planning to have an in-person gathering but dark clouds of COVID-19 continued to hover above us and forced us to make yet another difficult decision of hosting a virtual conference. While we are sure many of us were very much looking forward to an in-person gathering in Montreal, we welcome all participants to the virtual meeting. We hope this experience during the COVID19 pandemic can be as constructive as possible and lead to new insights and collaborations for the future. The asynchronous nature and recorded sessions allow all participants to view and review multiple times as much of the conference material as they please, something that is simply not possible with in-person meetings.

We have an exciting technical program enriched by an outstanding set of 9 Plenary Speakers from both academia and industry. Our plenary speakers are Prof. Ioannis Pitas (FIEEE, EURASIP fellow), Prof. Imre Rudas (FIEEE, IEEE SMCS President); Prof. Robert Kozma (FIEEE); Prof. Yaochu Jin (FIEEE, IEEE CIS VP); Dr. Anthony Vetro (FIEEE, VP & Director MERL); Prof. Henry Leung (FIEEE, FSPIE); Prof. Carlo S. Regazzoni (Chair IEEE ASI); Dr. Ming Hou (Defence Research & Development Canada), and Prof. Hagit Messer (FIEEE), who will give live video presentations and Q&A sessions. Each of these keynote speakers is absolute world-leader in their research fields and without a doubt would provide interesting new insights to their current and past research, and also their perspective on the future of autonomous systems. In addition to the 9 plenary talks, the inaugural IEEE ICAS'21 features a very selective main track, complemented by 5 special tracks and 6 special sessions. The acceptance rate for regular papers in the main track is about 25% out of all submissions received, while that for special sessions is around 40%. IEEE ICAS'21 also features an invited Panel entitled “Future Development of Autonomous Systems,” with the goal of providing our audience an opportunity to hear expert perspectives on future trends in the development of autonomous systems. Last but not least, a special initiative has been introduced with the inclusion of AutoDefence Summer School session consisting of 8 Tutorials delivered by internationally well-known speakers. This session will provide a platform for up-and-coming researchers and will be an exciting venue for sharing novel and new ideas.

We are in gratitude to the ICAS 2021 team for the organization of the conference, especially the Technical Program Committee (TPC) chair and vice-chairs, Special Track chairs, AutoDefence Summer School organizers, TPC members, and Conference Catalyst, for their enormous efforts. We would like to thank everyone else who were involved in the planning and organization of the conference. Special thanks are due to the conference participants for their excellent contributions making this inaugural conference a dynamic and informative one under exceptionally difficult circumstances. A particular mention should also be made of our sponsors; please ensure you check their details on the conference website and show your support by contacting and discussing with them any current or future needs you may have associated with their products or services.

We wish everyone a great virtual IEEE ICAS Conference in 2021. Bienvenue dans le monde passionnant du système autonome!

General Chairs
Arash Mohammadi and Amir Asif

ICAS 2021 Organizing Committee

General Chairs
Arash Mohammadi
Amir Asif

Technical Program Chair
Yingxu Wang

Technical Program Vice-Chairs
Lucio Marcenaro
Farokh Atashzar

Special Sessions Chairs
Mark Coates
Malika Meghjani

Local Arrangements Chairs
Warren Gross
Farnoosh Naderkhani

Finance Chair
Fabrice Labeau

Publicity Chairs
Lina Karam
Marcelo Bruno

Publications Chairs
Svetlana Yanushkevich
Neda Nategh

Workshop Chair
Marina L. Gavrilova

Industry Liaison Chairs
Mohammad Faghani
Morgan Kernohan

Award Chair
Saif Zahir

Special Track 1 Chairs
Dario Farina
Mahdi Tavakoli

Special Track 2 Chairs
Deepa Kundur
Mohammad Al Janaideh

Special Track 3 Chairs
Giuseppe Franze
Walter Lucia

Special Track 4 Chairs
Chun Wang
Anjali Awashthi

Special Track 5 Chairs
Usman Khan
Hoi-To Wai

Sponsors

Conference Sponsors

Summer School Sponsor

Table of Contents

Parallel Session
MT1: Framework of Autonomous Systems
Session Chair: Yingxu Wang

Parallel Session
ST1: Autonomous Medical Robotic Systems
Session Chair: Mahdi Tavakoli, Farokh Atashzar, and Dario Farina

Parallel Session
MT2: Emerging Technologies for Autonomous Systems (I)
Session Chair: Lucio Marcenaro

Parallel Session
SS1: Autonomous Vehicle Vision
Session Chair: Rui Fan and Ioannis Pitas

Parallel Session
SS2: Advanced Navigation for Networked AS'
Session Chair: Siwei Zhang and Henk Wymeersch

Parallel Session
ST2: Security & Resilience of Auto. Cyber-Physical Sys.
Session Chair: Deepa Kundur and Mohammad Janaideh

Parallel Session
MT3: Autonomous System and AI
Session Chair: Farokh S. Atashzar

Parallel Session
ST3: Autonomous Control Systems
Session Chair: Giuseppe Franze and Walter Lucia

Parallel Session
MT4: Application Paradigms of Autonomous Systems
Session Chair: Amir Asif

Parallel Session
SS3: Trustworthy Autonomous Human-Machine Systems
Session Chair: Yaoping Hu and Baris Fidan

Parallel Session
SS4: Explainable Machine Learning for AS'
Session Chair: Yong M. Ro, Parnian Afshar and Kostas Plataniotis

Parallel Session
ST4: Autonomous Transportation Systems
Session Chair: Chun Wang and Anjali Awasthi

Parallel Session
MT5: AS Solutions for Engineering Problems
Session Chair: Farokh Atashzar

Parallel Session
ST5: Signal Processing for Self-Aware & Social AS'
Session Chair: Hoi-To Wai and Usman Khan

Parallel Session
MT6: Emerging Technologies for Autonomous Systems (II)
Session Chair: Arash Mohammadi

Parallel Session
SS5: Autonomous Diagnosis/Prognosis of COVID-19
Session Chair: Farnoosh Naderkhani and Moazedin J. Rafie

Plenary Panel
Future Development of Autonomous Systems
Session Chair: Yingxu Wang

Drone Vision and Deep Learning for Infrastructure Inspection

Ioannis Pitas, *FIEEE, EU RASIP Fellow*

Director, Artificial Intelligence and Information Analysis (AIIA) Lab.
Department of Informatics
Aristotle University of Thessaloniki (AUTH), Greece
Email: *pitas@csd.auth.gr*

ABSTRACT

This lecture overviews the use of drones for infrastructure inspection and maintenance. Various types of inspection, e.g., using visual cameras, LIDAR or thermal cameras are reviewed. Drone vision plays a pivotal role in drone perception/control for infrastructure inspection and maintenance, because: a) it enhances flight safety by drone localization/mapping, obstacle detection and emergency landing detection; b) performs quality visual data acquisition, and c) allows powerful drone/human interactions, e.g., through automatic event detection and gesture control. The drone should have: a) increased multiple drone decisional autonomy and b) improved multiple drone robustness and safety mechanisms (e.g., communication robustness/safety, embedded flight regulation compliance, enhanced crowd avoidance and emergency landing mechanisms). Therefore, it must be contextually aware and adaptive. Drone vision and machine learning play a very important role towards this end, covering the following topics: a) semantic world mapping b) drone and target localization, c) drone visual analysis for target/obstacle/crowd/point of interest detection, d) 2D/3D target tracking. Finally, embedded on-drone vision (e.g., tracking) and machine learning algorithms are extremely important, as they facilitate drone autonomy, e.g., in communication-denied environments. Primary application area is electric line inspection. Line detection and tracking and drone perching are examined. Human action recognition and co-working assistance are overviewed.

The lecture will offer: a) an overview of all the above plus other related topics and will stress the related algorithmic aspects, such as: b) drone localization and world mapping, c) target detection d) target tracking and 3D localization e) gesture control and co-working with humans. Some issues on embedded CNN and fast convolution computing will be overviewed as well.

About the Keynote Speaker

Prof. Ioannis Pitas (IEEE fellow, IEEE Distinguished Lecturer, EURASIP fellow) received the Diploma and PhD degree in Electrical Engineering, both from the Aristotle University of Thessaloniki (AUTH), Greece. Since 1994, he has been a Professor at the Department of Informatics of AUTH and Director of the Artificial Intelligence and Information Analysis (AIIA) lab. He served as a Visiting Professor at several Universities. His current interests are in the areas of computer vision, machine learning, autonomous systems, intelligent digital media, image/video processing, human-centred computing, affective computing, 3D imaging and biomedical imaging. He has published over 1000 papers, contributed in 47 books in his areas of interest and edited or (co-)authored another 11 books. He has also been member of the program committee of many scientific conferences and workshops. In the past he served as Associate Editor or co-Editor of 9 international journals and General or Technical Chair of 4 international conferences. He participated in 71 R&D projects, primarily funded by the European Union and is/was principal investigator in 42 such projects. Prof. Pitas leads the big European H2020 R&D project MULTIDRONE: https://multidrone.eu/. He is AUTH principal investigator in H2020 R&D projects Aerial Core and AI4Media. He is chair of the Autonomous Systems Initiative https://ieeeasi.signalprocessingsociety.org/. He is head of the EC funded AI doctoral school of Horizon2020 EU funded R&D project AI4Media (1 of the 4 in Europe). He has 32000+ citations to his work and h-index 85+ (Google Scholar).

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

VERIFICATION, TRUSTWORTHINESS AND ACCOUNTABILITY OF HUMAN-DRIVEN AUTONOMOUS SYSTEMS

Imre Rudas and Tamas Haidegger*

University Research and Innovation Center (EKIK), Óbuda University, Budapest, Hungary
* rudas@uni-obuda.hu

ABSTRACT

Despite the fact that autonomous systems' science and control theory have almost 50 years of history, the community is facing major challenges to ensure the safety of fully autonomous consumer systems. It mostly concerns the verification and high fidelity operation of safety-critical systems, may that be a self-driving car, a homecare robot or a surgical manipulator. The community still struggles to establish objective criteria for trustworthiness of AI driven / machine learning based control systems. On one hand, we celebrate the rise of cognitive capabilities in robotic systems, leading independent decision making; on the other hand, decisions made in complex environments, based on multi-sensory data will surly lead to some wrong conclusions and hazardous outcome, jeopardizing the public trust in entire application domains. This ambiguity led to the currently ruling safety principle to offer the possibility for a human-driven override, translating to Level of Autonomy 3 and 4 with autonomous vehicles.

The aim of the development community is to establish processes and metrics to ensure the reliability of the takeover process, when the human driver or operator takes back the partial or full control from the autonomous system. We have been building complex simulators and data collection systems to benchmark human decision making against the computer. Situation Awareness (SA) has been identified as a key, as it defines the level of cognitive understanding and capability of a human operator in a given environment. Assessing, maintaining and regaining efficiently SA are core elements of the relevant research projects, reviewed and compared in this talk. Based on the research at the Antal Bejczy Center for Intelligent Robotics at Óbuda University, we created an assessment method for critical handover performance, to quantitatively define the required level and components of SA with respect to the autonomous functionalities present. To improve system safety, driver assistance systems and automated driving functionalities shall be collected and organized in a hierarchical way, along with the two criteria of SA presented, as a standardized risk assessment protocol: 1) the Level of SA, based on state of the environment; 2) the components of SA, based on knowledge.

The outcome of our experiments may find its way to new verification standards through ongoing IEEE initiatives, such as the P1872.1, P2817, P7000 and P7007, moreover this systematic approach has already proved to bring benefit to other domains, such as medical robotics.

Keywords*: autonomous vehicle safety; situation awareness; Level of Autonomy; human takeover; hands-off control*

Imre J. Rudas (M'91-F'02-LM'20) received the graduation degree from Banki Donat Polytechnic, Budapest, Hungary in 1971, received the Master Degree in mathematics from the Eotvos Lorand University, Budapest, the Ph.D. in robotics from the Hungarian Academy of Sciences in 1987, while the Doctor of Science degree from the Hungarian Academy of Sciences in 2004. He is Professor Emeritus of Óbuda University and Advisory Board Chair of the University Research and Innovation Center. He has published 6 books, 12 university books, more than 850 papers in various journals and international conference proceedings, and received more than 7000 citations. His present areas of research activities are Computational Cybernetics, Robotics, Computational Intelligence. Prof. Rudas is the President of the IEEE Systems, Man, and Cybernetics Society.

Tamás Haidegger (M'04-SM'18) received his MSc degrees from the Budapest Uni. of Technology and Economics in EE and BioE, then PhD in medical robotics. His main field of research is on autonomous system, control/teleoperation of surgical robots, image-guided therapy and digital health technologies. Currently, he is associate professor at Óbuda University, serving as the director of EKIK. He is an active member of the IEEE Robotics and Automation Society (serving as an associate VP), IEEE SMC, IEEE EMBC, IEEE SA, holding leadership positions in the IEEE Hungary Section as well. Tamas is the author and co-author of over 250 scientific publications, with over 2000 independent citations. He has a professional blog: *surgrob.blogspot.com*.

Sustainable Autonomy of Intelligent Systems: Challenges and Perspectives

Robert Kozma, *FIEEE*
Department of Mathematics, FedEx Institute of Technology
The University of Memphis, Memphis TN 38120, USA
Email: *rkozma@memphis.edu*

Abstract

Cutting-edge autonomous systems demonstrate outstanding performance in many important tasks requiring intelligent data processing under well-known conditions, supported by massive computational resources and big data. However, the performance of these systems may drastically deteriorate when the data are perturbed, or the environment dynamically changes, either due to natural effects or caused by man-made disturbances. The challenges are especially daunting in edge computing scenarios and on-board applications with limited resources, due to constraints on the available data, energy, computational power, while critical decisions must be made rapidly, in a robust way. A neuromorphic perspective provides crucial support under such conditions. Human brains are efficient devices using 20W power (just like a light bulb!), which is drastically less than the power consumption of today's supercomputers requiring MWs to solve specific learning tasks in an innovative way. This is not sustainable. Brains use spatio-temporal oscillations to implement pattern-based computing, going beyond the sequential symbol manipulation paradigm of traditional Turing machines. Neuromorphic spiking chips, including memristor technology, provide crucial support to the field. Application examples include on-board signal processing, distributed sensor systems, autonomous robot navigation and control, and rapid response to emergencies.

Keywords: *Autonomous Systems, Edge Computing, Neuromorphic Computing; Brain Computing; Sustainable AI*

I. Resource Constraints of State-of-Art Intelligent Systems Developments

The past decades demonstrated massive proliferation of intelligent systems based on powerful AI technologies in a wide range of applications, including manufacturing, health care, transportation education, and finances. The dominant approaches in these applications use Deep Learning and produce cutting-edge AI with often super-human performance [1]–[3]. In spite of their enormous successes, AI implementations face challenges due to their rigidity, lack of robustness, and difficulty to adopt to changing conditions [4]. Deep Learning poses high demands on computational resources and it requires the collection and maintenance of huge data resources, which represent serious challenges both from engineering, and societal, ethical perspectives [**?**]. There is a prominent view that we are rapidly approaching the end of Moore's law [5]. In fact, the end or Moore's law may have arrived, demanding a drastic reformulation of digital technology dominating computer hardware developments for over half century [6].

A key impediment of computer hardware development is the heat produced in the densely packed electronics on microchips. The problem is twofold: (1) the heat must to be removed, which is increasingly difficult due to the miniaturization of electronics components; (2) the energy dissipated in the form of heat needs to be produced, which is a nontrivial matter, as today's advanced Deep Learning computers operate using many MWs of electrical power. These computers can solve problems at human-level, or even exceeding human performance. Still, it worth to note that human brains use 20W of power, just as a lightbulb, thus brains need millions of times less energy than supercomputers for AI tasks. With the proliferation of Deep Learning, the trend of using massive energy resources is accelerating. Figure 1. illustrates the exponentially increasing computational demand in the last few years, from Alexnet to AlphaGo Zero [7]. There are various attempts to address the increasing energy demand, e.g., by making our chips more energy-efficient, but these approaches fall short of addressing the fundamental problems ahead and new solutions are needed [6], [8].

Situated intelligence can supplement traditional AI to mitigate some of the mentioned shortcomings due to inherent resource constraints. Situated intelligence views intelligent systems embodied in their environment as they develop solutions to their tasks [9]–[11]. Embodiment imposes constraints on the system's intelligence as it evolves, and energy constraints are important aspects of embodiment. Considering energy constraints, intelligent systems can develop solutions which are inherently energy efficient. In this chapter, we review complimentary aspects of intelligence and provide an example how energy constrains can be incorporated to intelligent systems designs. In this approach, we benefit from lessons learnt from biological brains combining symbolic computation and subsymbolic metabolic processing [12], [13]. The goal is to mitigate the problems of inefficient resource utilization in mainstream AI solutions and to provide a perspective for a sustainable AI development in the decades ahead. Brain dynamics provide important clues for the development of artificially intelligent systems with efficient use of energy. We

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

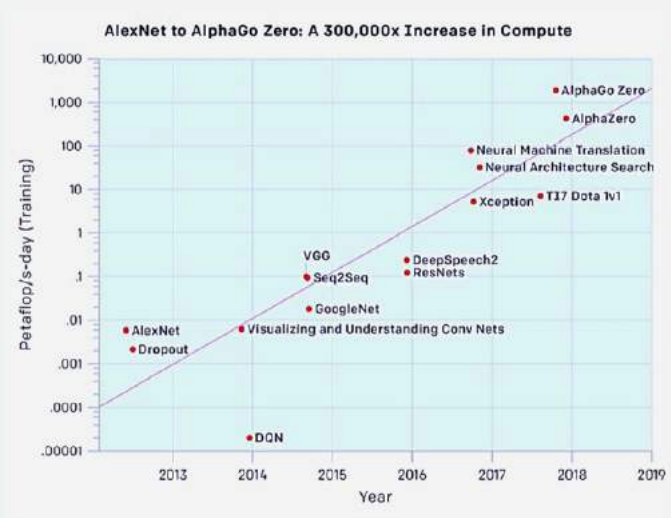


Fig. 1. Illustration of the exponentially increasing computational demand of Deep learning to solve various machine learning tasks during the past decade; adopted from OpenAI; adopted from [7]

Computing Features	Brain Dynamics	Pattern-based Computing	Turing Machine
Domains of Realization	Human Neocortex	Neuromorphic Chip	Digital Supercomputer
Power consumed in challenging Machine Learning task	20W	$20x10^3W$*	$20x10^6W$**
Units of computing employed	Cortical AM patterns	Computational AM patterns	Fixed predefined symbols
Rules of computing executed	Implicit rules manifested at theta rate (5Hz)	Phase transitions between AM patterns	Fixed preset rule base

Fig. 2. Comparison of Computational Approaches: Brains, Neuromorphic Chips, and Turing Machines. Notations: ($\star$) Value estimated based on computing with dedicated neuromorphic hardware platforms [8]; ($\star\star$) Estimation based on AlphaGo power use of 1MW in 2016, and a 20-fold increase of the required increase of compute resources since then, as shown in Fig.1.

learn from the hierarchy of neural structures, dynamics, and corresponding functions, to achieve this goal by neuromorphic computing approaches.

II. Neuromorphic Intelligent Systems with Brain-like Dynamics for Efficient Operation

Brain dynamics provides important clues for the development of artificially intelligent systems with efficient use of energy. We emphasize here two aspects of brain operation, such as (i) pattern based computing and (ii) neuromorphic computing with spiking neural networks, in order to develop intelligent systems that can realize sustainable utilization of resources, including energy.

- *Pattern-based computing* involves the hierarchy of neural structures, dynamics, and functions. At the top of the hierarchy there is the macroscopic level, corresponding the whole brain, while the microscopic cellular level represents the opposite end. These two opposing aspects are connected through mesoscopic levels, which are critical in maintaining a balance across vast scales of hierarchy [14], [15]. Brains are intelligent systems which reconcile the opposing aspects of local fragmentation and global uniform dominance of a single state. These opposing aspects coexist in the brain in the form of metastable activity patterns. Pattern-based computing principles in intelligent systems designs utilize such activity patterns and they go beyond the computing paradigm based on Turing machines dominating the digital computing landscape for over 70 years [16], [17]. The patterns used in this brain-like computing are emergent properties of the oscillating media, thus they require drastically less resources to be maintained [18], [19].
- *Neuromorphic technologies* have been intensively developed in recent years using novel hardware designs with spiking neural networks [8], [20], [21]. Such hardware platforms can be used in large-scale networks with recurrent connections. These platforms are very powerful, but they often lack the capability of adaptation and learning in spiking domain, which are important requirements for intelligent systems applications. In recent years, very efficient spiking neural network simulator platforms have been developed, including spiking time-dependent plasticity (STDP), which were employed using cutting-edge neuromorphic computing platforms [22]–[26]. Neuromorphic technologies achieve significantly improved power utilization compared to leading state-of-art AI approaches; see Figure 2. Figure 2 compares various attributes of brain computing, Turing machines, and pattern-based computing. Human brains use approx. 20W power, while today's cutting-edge AlphaZero hardware demands 20MW to solve a challenging machine learning task; see Fig. 1. For some specific AI tasks, DL uses a million times more energy than brains. Neuromorphic hardware with crossbar architecture of spiking neurons is much more efficient than digital supercomputers, still it uses 1,000 times more energy than brains.

III. Conclusion

The development of increasingly powerful intelligent systems in the past decades has been based on exponentially advancing digital computing technology. This extensive development is not sustainable in the foreseeable future and it demands the drastic reformulation of existing approaches. A key issue is the massive energy utilization by the electronics components. Energy constraints are often ignored or have just secondary role in typical cutting-edge AI approaches. Based on lessons learnt from neuroscience and brain dynamics, new generation of brain-inspired intelligent systems are developed which use energy more efficiently that mainstream technology. Neuromorphic hardware systems combined with pattern-based computing principles produce intelligent systems with the desirable properties for sustainable intelligent systems.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[2] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[3] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[4] G. Marcus, "Deep learning: A critical appraisal," *arXiv preprint arXiv:1801.00631*, 2018.

[5] R. S. Williams, "What's next?[the end of moore's law]," *Computing in Science & Engineering*, vol. 19, no. 2, pp. 7–13, 2017.

[6] M. M. Waldrop, "The chips are down for moore's law," *Nature News*, vol. 530, no. 7589, p. 144, 2016.

[7] D. Amodei and H. D., "Ai and compute," *https://openai.com/blog/ai-and-compute/ blog*, vol. 8/19/2019, 2018.

[8] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *Ieee Micro*, vol. 38, no. 1, pp. 82–99, 2018.

[9] R. Brooks, *Cambrian intelligence: The early history of the new AI*. Cambridge, MA, USA: MIT Press, 1999, vol. 97.

[10] H. L. Dreyfus, "Why heideggerian ai failed and how fixing it would require making it more heideggerian," *Philosophical psychology*, vol. 20, no. 2, pp. 247–268, 2007.

[11] R. Kozma, H. Aghazarian, T. Huntsberger, E. Tunstel, and W. J. Freeman, "Computational aspects of cognition and consciousness in intelligent devices," *IEEE Computational Intelligence Magazine*, vol. 2, no. 3, pp. 53–64, 2007.

[12] R. Noack, C. Manjesh, M. Ruszinko, H. Siegelmann, and R. Kozma, "Resting state neural networks and energy metabolism," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 228–235.

[13] R. Kozma, "Computers versus brains: Game is over or more to come?" in *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Elsevier, 2019, pp. 205–218.

[14] W. J. Freeman, "The physiology of perception," *Scientific American*, vol. 264, no. 2, pp. 78–87, 1991.

[15] R. Kozma and W. J. Freeman, *Cognitive phase transitions in the cerebral cortex-enhancing the neuron doctrine by modeling neural fields*. Springer, 2016.

[16] A. Turing, "Machinery and intelligence," *Mind: A Quarterly Review of Psychology and Philosophy*, vol. 59, p. 236, 2003.

[17] J. Von Neumann, *The computer and the brain*. Yale University Press, 1956.

[18] R. Kozma and M. Puljic, "Hierarchical random cellular neural networks for system-level brain-like signal processing," *Neural Networks*, vol. 45, pp. 101–110, 2013.

[19] ——, "Pattern-based computing via sequential phase transitions in hierarchical mean field neuropercolation," *Theoretical Computer Science*, vol. 633, pp. 54–70, 2016.

[20] M. V. DeBole, B. Taba, A. Amir, F. Akopyan, A. Andreopoulos, W. P. Risk, J. Kusnitz, C. O. Otero, T. K. Nayak, R. Appuswamy *et al.*, "Truenorth: Accelerating from zero to 64 million neurons in 10 years," *Computer*, vol. 52, no. 5, pp. 20–29, 2019.

[21] S. J. Van Albada, A. G. Rowley, J. Senk, M. Hopkins, M. Schmidt, A. B. Stokes, D. R. Lester, M. Diesmann, and S. B. Furber, "Performance comparison of the digital neuromorphic hardware spinnaker and the neural network simulation software nest for a full-scale cortical microcircuit model," *Frontiers in neuroscience*, vol. 12, p. 291, 2018.

[22] S. Furber, "Large-scale neuromorphic computing systems," *Journal of neural engineering*, vol. 13, no. 5, p. 051001, 2016.

[23] K. Roy, A. Jaiswal, and P. Panda, "Towards spike-based machine intelligence with neuromorphic computing," *Nature*, vol. 575, no. 7784, pp. 607–617, 2019.

[24] H. Hazan, D. J. Saunders, H. Khan, D. Patel, D. T. Sanghavi, H. T. Siegelmann, and R. Kozma, "Bindsnet: A machine learning-oriented spiking neural networks library in python," *Frontiers in neuroinformatics*, vol. 12, p. 89, 2018.

[25] D. Patel, H. Hazan, D. J. Saunders, H. T. Siegelmann, and R. Kozma, "Improved robustness of reinforcement learning policies upon conversion to spiking neuronal network platforms applied to atari breakout game," *Neural Networks*, vol. 120, pp. 108–115, 2019.

[26] D. J. Saunders, D. Patel, H. Hazan, H. T. Siegelmann, and R. Kozma, "Locally connected spiking neural networks for unsupervised feature learning," *Neural Networks*, vol. 119, pp. 332–340, 2019.

Dr. Robert Kozma holds a Ph.D. in Applied Physics (Delft University of Technology, The Netherlands, 1992). He has two M.Sc. degrees, one in Applied Mathematics with honors (Roland Eötvös University, Budapest, Hungary, 1988, another one in Physical Engineering with distinction (Moscow Institute of Power Engineering MEI, Moscow, Soviet Union, 1982).

He has been Professor of Mathematics and Computer Science at the University of Memphis, TN, USA since 2000, where he is funding Director of Center for Large-Scale Intelligent Optimization and Networks (CLION), FedEx Institute of Technology. Visiting positions include Professor of Computer Science, University of Massachusetts Amherst, MA; US Air Force Research Laboratory, Sensors Directorate, WPAFB, OH; NASA Jet Propulsion Laboratory, Robotics; Caltech, Pasadena, CA. Previous affiliations include University of California at Berkeley, EECS and Div. Neurobiology (1998-2000); Otago University, Information Sciences, New Zealand (1996-1998); Tohoku University, Quantum Science and Engineering, Japan (1993-1996). He has over 35 years of experience in intelligent signal processing, anomaly detection, autonomous systems, large-scale networks, distributed sensor systems, and biomedical applications. Published 9 books/edited volumes, over 300 papers, has 3 patents. Gave over 200 talks at conferences, about half of them are plenary, keynote, and invited talks. Extensive research support include NASA on cognitive robotics; Defense Advanced Research Projects Agency (DARPA) on physical intelligence and energy-aware superior AI; Air Force Research Laboratory (AFRL), Sensors Directorate on distributed sensing; Air Force Office on Scientific Research (AFOSR) on neurodynamics of perception and cognition; National Science Foundation (NSF) on strategy changes in cognition and technology, and graph theory for brain networks; Office of Naval Research (ONR) on nonlinear brain dynamics and attractor networks; and other agencies.

Dr. Robert Kozma is Fellow of the IEEE and Fellow of the International Neural Networks Society (INNS). He has been President of INNS (2017-2018), served on the Board of Governors of IEEE Systems, Man and Cybernetics Society (2016-2018, 2020), AdCom of IEEE Computational Intelligence Society (2009- 2012), and INNS Board of Governors (2007-2012).

He has been General Chair of IJCNN2009, and served as Program Chair/Co-Chair of dozens of conferences. He is Editor-In-Chief of IEEE Transactions of Systems, Man, and Cybernetics - Systems. He has been Associate Editor, including Neural Networks, IEEE Transactions on Cybernetics, IEEE Transactions on Neural Networks, Cognitive Systems Research, Cognitive Neurodynamics. He is recipient of the INNS Dennis Gabor Award.

Morphogenetic Self-Organization of Collective Systems

Yaochu Jin, *Fellow IEEE*
Department of Computer Science, University of Surrey, UK
Email: *yaochu.jin@surrey.ac.uk*

Abstract

Self-organization is one of the most important features observed in social, economic, ecological and biological systems. Distributed self-organizing systems are able to generate emergent global behaviors through local interactions between individuals without a centralized control. Such systems are supposed to be robust, self-repairable and highly adaptive. However, design of self-organizing systems is very challenging, particularly when the emerged global behaviors are required to be predictable or predictable. This talk introduces a morphogenetic approach to the self-organizing swarm robots using genetic and cellular mechanisms governing the biological morphogenesis. We demonstrate that morphogenetic self-organizing algorithms are able to autonomously generate patterns and surround moving targets without centralized control. Finally, morphogen based methods for self-organization of simplistic robots that do not have localization and orientation capabilities are presented.

Keywords: *Morphogenetic development, diffusion-reaction, self-organizing systems, swarm robots*

About the Keynote Speaker

Yaochu Jin received the BSc, MSc, and PhD degrees from Zhejiang University, Hangzhou, China, in 1988, 1991, and 1996, respectively, and the Dr.-Ing. degree from Ruhr University Bochum, Germany, in 2001.

He is currently a Distinguished Chair, Professor in Computational Intelligence, Department of Computer Science, University of Surrey, Guildford, U.K., where he heads the Nature Inspired Computing and Engineering Group. He was a "Finland Distinguished Professor" of University of Jyvaskyla, Finland, a "Changjiang Distinguished Visiting Professor", Northeastern University, China, and "Distinguished Visiting Scholar", University of Technology Sydney, Australia. His main research interests include data-driven surrogate-assisted evolutionary optimization, trustworthy machine learning, multi-objective evolutionary learning, swarm robotics, and evolutionary developmental systems.

Dr Jin is presently the Editor-in-Chief of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS and the Editor-in-Chief of Complex & Intelligent Systems. He was an IEEE Distinguished Lecturer, and Vice President of the IEEE Computational Intelligence Society. He is the recipient of the 2018 and 2020 IEEE Transactions on Evolutionary Computation Outstanding Paper Award, the 2014, 2016, and 2019 IEEE Computational Intelligence Magazine Outstanding Paper Award, and the Best Paper Award of the 2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology. He is recognized as a Highly Cited Researcher 2019 and 2020 by the Web of Science Group. He is a Fellow of IEEE.

Dr Jin was recently awarded the Alexander von Humboldt Professorship for Artificial Intelligence by the German Federal Ministry of Education and Research, and is going to move to the Bielefeld University in October 2021.

IMPROVING MANIPULATION CAPABILITIES OF AUTONOMOUS ROBOTS

Anthony Vetro

Mitsubishi Electric Research Labs, Cambridge, MA, USA
Email: avetro@merl.com

ABSTRACT

Human-level manipulation continues to be beyond the capabilities of today's robotic systems. Not only do current industrial robots require significant time to program a specific task, but they lack the flexibility to generalize to other tasks and be robust to changes in the environment. While collaborative robots help to reduce programming effort and improve the user interface, they still fall short on generalization and robustness. This talk will highlight recent advances in a number of key areas to improve the manipulation capabilities of autonomous robots, including methods to accurately model the dynamics of the robot and contact forces, sensors and signal processing algorithms to provide improved perception, optimization-based decision-making and control techniques, as well as new methods of interactivity to accelerate and enhance robot learning.

Index Terms— robotics, learning, manipulation, generalization, robustness, perception, control.

PLENARY SPEAKER BIO

Anthony Vetro (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from Polytechnic University, Brooklyn, NY.

Dr. Vetro is currently VP & Director at Mitsubishi Electric Research Labs, in Cambridge, MA. He is responsible for AI related research in the areas of computer vision, speech/audio processing, and data analytics. In his 25 years with the company, he has contributed to the development and transfer of several technologies to Mitsubishi products, including digital television receivers and displays, surveillance and camera monitoring systems, automotive equipment, as well as satellite imaging systems. He has published more than 200 papers and has been a member of the MPEG and ITU-T video coding standardization committees for a number of years, serving in numerous leadership roles.

Dr. Vetro is active in various IEEE conferences, technical committees and editorial boards. Since 2019, he has been serving as a Senior Associate Editor of the IEEE Open Journal on Signal Processing. Past roles include serving on the SPS Conference Board and the Conference Board Executive Subcommittee (2018-2019), and the SPS Nominations Appointments Committee (2018-2019). He also served as a Senior Editorial Board member of the IEEE Journal on Selected Topics in Signal Processing (2013-2015) and IEEE Journal on Emerging and Selected Topics in Circuits and Systems (2016-2017); on the Editorial Board of IEEE Signal Processing Magazine (2009-2011) and IEEE Multimedia (2010-2016); as an Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology (2010-2013), IEEE Transactions on Image Processing (2010-2014) and APSIPA Transactions on Signal and Information Processing (2012-2017); and on the Steering Committee of IEEE Transactions on Multimedia (2008-2010). In the IEEE Signal Processing Society, he served as Chair of the TC on Multimedia Signal Processing of (2008-2009) and as a Member of the Image, Video and Multidimensional Signal Processing (2010-2012). He was a General Co-Chair of ICIP 2017 in Beijing, ICME 2015 in Torino and MMSP 2011 in Hangzhou, and also served as a Technical Program Co-Chair for ICME 2016 in Seattle. He also served on the IEEE Fellows Evaluation Committee of the IEEE Circuits Systems Society (2011-2017). He has received several awards for his work on transcoding and is a Fellow of the IEEE.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Information Fusion and Decision Support for Autonomous Systems

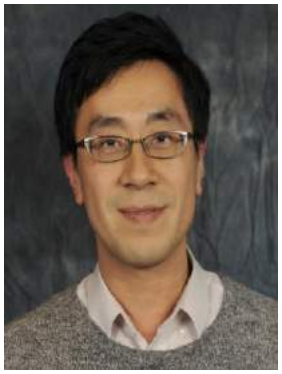
Henry Leung

Dept. Electrical and Software Engineering
University of Calgary
2500 University Drive NW, Calgary, AB, Canada T2N 1N4
Email: *leungh@ucalgary.ca*

ABSTRACT

In this talk we present our works on decision support analytic for autonomous systems. Decision support analytic process multiple sensory information collected by an autonomous system such as lidar, camera, RGBD, acoustic to perform signal detection, target tracking, object recognition. As multiple sensors are involved, our system uses sensor registration, data association and fusion to combine sensory information. The next layer of the proposed decision support system orients the processed sensory information at feature and classification levels to perform situation assessment and treat evaluation. Based on the assessment, the decision support system will recommend decision. If the uncertainty is high, actions including resource allocation, planning will be used to extract or reassess the sensory information to get a recommended decision with lower uncertainty. This talk will also presents the applications of the proposed decision support analytic in four industrial projects including 1) goal-driven net-enabled distributed sensing for maritime surveillance, 2) autonomous navigation and perception of humanoid service robots, 3) distance learning for oil and gas drilling and 4) cognitive vehicles.

About the Keynote Speaker



Henry Leung is a professor of the Department of Electrical and Computer Engineering of the University of Calgary. Before joining U of C, he was with the Department of National Defence (DND) of Canada as a defence scientist. His main duty there was to conduct research and development of automated surveillance systems, which can perform detection, tracking, identification and data fusion automatically as a decision aid for military operators. His current research interests include big data analytic, chaos and nonlinear dynamics, information fusion, machine learning, signal and image processing, robotics and internet of things. He has published extensively in the open literature on these topics. He has over 300 journal papers and 250 refereed conference papers. Dr. Leung has been the topic editor on "Robotic Sensors" of the International Journal of Advanced Robotic Systems and the associate editor of various journals such as the IEEE Circuits and Systems Magazine, International Journal on Information Fusion, Sensor Journal, IEICE Trans. Nonlinear Theory and its Applications, Aerospace Systems, IEEE Trans. Aerospace and Electronic Systems, IEEE Signal Processing Letters, IEEE Trans. Circuits and Systems. He has also served as guest editors for various special issues such as "Intelligent Transportation Systems" for the International Journal on Information Fusion and "Cognitive Sensor Networks" for the IEEE Sensor Journal. He is the editor of the Springer book series on "Information Fusion and Data Science". He is a Fellow of IEEE and SPIE.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Bayesian emergent self awareness

Carlo S. Regazzoni
Electrical, Electronics and Telecommunication Engineering and Naval Architecture Department (DITEN)
University of Genova
Genova, Italy
E.mail: carlo.regazzoni@unige.it

***Abstract*—Multisensor signal Data Fusion and Perception, including processing of signals are important cognitive functionalities that can be included in artificial systems to increase their level of autonomy. However, the techniques they rely on have been developed incrementally along time with the underlying assumption that they should have been used mainly to provide a support to decision tasks driving the actions of those systems. Cognitive functionalities like self-awareness have been so far considered as not primary part of embodied knowledge of an autonomous or semi autonomous systems. One of the reason for this choice was the lack of understanding the principles that could allow an agent, even a human one, to organize successive sensorial experiences into a coherent framework of emergent knowledge, by means of integrating signal processing, machine learning and data fusion aspects. However, the developments of this last decade in many fields carried to the possibility to provide integrated solutions capable to sketch how emergent self awareness can be obtained by capturing experiences of autonomous agents like for example vehicles and intelligent radios. In this presentation, a hierarchical Bayesian representation is proposed based on generalized random states and including in a coherent inference framework anomaly detection and incremental learning. Described models are provided of generative (temporally and hierarchically) predictive as well as of discriminative capabilities and can be used as bricks of emergent self awareness in intelligent agents. Discussion of the advantages of including emergent self awareness in intelligent agents will be also provided with respect to different aspects, e.g. explainability of agent's actions and capability of imitation learning.**

***Index Terms*—Self awareness; Dynamic Bayesian Networks; anomaly detection; incremental learning; imitation learning; generalized coordinates; generative models; discriminative models.**

About the Keynote Speaker

Carlo S. Regazzoni obtained the M.S. and PhD degrees from University of Genova, in 1987 and 1992, respectively. Since 2005, he is full professor of Cognitive Telecommunications systems at DITEN, University of Genova, Italy. He is coordinating international Interactive and Cognitive Environment PhD courses at UNIGE since 2008. His research interests include cognitive dynamic systems, adaptive and self-aware multimodal signal processing, Bayesian machine learning, Cognitive radio. He is author of peer-reviewed papers on more than 100 international journals and 350 at international conferences. He served in IEEE Signal Processing Society in many roles, including VP conferences in 2015-2017, Italy SPS Chapter Chair, 2010-2012, IEEE AVSS SC chair 2000-2010. He was General Chair, Technical Program chair and other roles in several international IEEE conferences within his research field. He is currently Chair of the IEEE SPS Autonomous Systems initiative and of the Fourier award board. He is/has been associate/guest editor of several int. journals including July 2020 special issue of the Proceedings of the IEEE on Self Aawraeness in Autonomous Systems.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

On the Ethics of Autonomous and Intelligent Systems (AIS)

Hagit Messer, Life Fellow of the IEEE
School of Electrical Engineering,
Tel Aviv University, Israel
Email: messer@eng.tau.ac.il

Abstract

In the 4th industrial revolution under which autonomous, intelligent systems are designed, certain human brain capacities are delegated to machines. This brings in great opportunity to reduce the need for human intervention in routines of daily lives, together with considerable ethical challenges. The role of the designers of such systems, i.e., engineers, is most important in balancing the opportunities and the challenges. Being a global organization with more than 450,000 members, IEEE took responsibility and has set up the Global Initiative on Ethics of Autonomous and Intelligent Systems, whose mission is: "To ensure every stakeholder involved in the design and development of autonomous and intelligent systems is educated, trained, and empowered to prioritize ethical considerations so that these technologies are advanced for the benefit of humanity." In this talk I will present the various activities of the global initiative to promote ethics in AIS and, in particular, I will introduce *Ethically Aligned Design*, *First Edition*, which is a comprehensive treatise that combines a conceptual framework addressing universal human values, sustainability, data agency, and technical dependability, among other pressing issues with a set of principles to guide A/IS creators and users through a wide-ranging set of recommendations, further resources and reference material, and a painstakingly created glossary.

Moreover, users - at all levels - of artificial intelligent systems must be aware of the ethical implications when deciding to use them. Policy makers are responsible for their regulation. But most importantly, designers of AIS systems - engineers - who are the only ones that can control some of the features reflecting on ethical issues (e.g., transparency) must take personal responsibility for the systems they produces to make sure they prioritize ethical considerations so that these technologies are advanced for the benefit of humanity. The way to raise awareness and responsibility is education, and IEEE has a special role in educating present and future engineers in the ethically aligned design of autonomous, intelligent systems.

Keywords: *AI ethics, Ethically Aligned Design, Cognitive Algorithms.*

About the Keynote Speaker

HAGIT MESSER received the Ph.D. in Electrical Engineering from Tel Aviv University (TAU), ISRAEL, and after a post-doctoral fellowship at Yale University, she joined the faculty of Engineering at Tel Aviv University in 1986, where she is The Kranzberg Chair Professor in Signal Processing at the school of Electrical Engineering. On 2000-3 she has been on leave from TAU, serving as the ***Chief Scientist*** at the Ministry of Science. After returning to TAU she was the head of the Porter school of environmental studies (2004-6), and the ***Vice President for Research and Development*** (2006-8). Then, she has been the ***President*** of the Open University, Israel (2008-13), and from Oct. 2013 till January 2016 she has served as the Vice Chair of the Council of Higher Education, Israel.

She was also one of the co-founders of the start-up company ClimaCell. Prof. Messer, Life Fellow of the IEEE, is an expert in statistical signal processing with applications to source localization, communication and environmental monitoring. She has published numerous journal and conference papers, and has supervised more than hundred graduate students. She has served as a member of Technical committees of the Signal Processing society since 1993 and on the editorial boards of the IEEE Transactions on Signal Processing, the IEEE Signal Processing Letters, the IEEE journal of selected topics in signal processing (J-STSP), and on the Overview Editorial Board of the Signal Processing Society journals.

Prof. Messer is interested in various aspects of higher education and science policy, including women in science and technology, and commercialization of academic research. Her interests, technological background and her acquaintance with governance systems have led her to get involved with various ethical committees. Since 2015 she is a member of COMEST, UNESCO's committee for ethics in Science and Technology, under which she was part of a working group which was responsible for a report on ethics in robotics, and then ethics of AI and recently ethics of IoT. On 2016 she has joined the Executive Committee of The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (AIS), where she is the chair of its Education Subcommittee. Since 2020 she is also a member of IEEE Ethics and Member Conduct Committee (EMCC).

Enabling Trust in Autonomous Human-Machine Teaming

Dr. Ming Hou
Defence Research & Development Canada
Email: *Ming.Hou@forces.gc.ca*

Abstract

The advancement of AI enables the evolution of machines from relatively simple automation to completely autonomous systems that augment human capabilities with improved quality and productivity in work and life. The singularity is near! However, humans are still vulnerable. The COVID-19 pandemic reminds us of our limited knowledge about nature. The recent accidents involving Boeing 737 Max passengers ring the alarm again about the potential risks when using human-autonomy symbiosis technologies. A key challenge of safe and effective human-autonomy teaming is enabling "trust" between the human-machine team. It is even more challenging when we are facing insufficient data, incomplete information, indeterministic conditions, and inexhaustive solutions for uncertain actions. This calls for the imperative needs of appropriate design guidance and scientific methodologies for developing safety-critical autonomous systems and AI functions. The question is how to build and maintain a safe, effective, and trusted partnership between humans and autonomous systems. This talk discusses a context-based and interaction-centred design (ICD) approach for developing a safe and collaborative partnership between humans and technology by optimizing the interaction between human intelligence and AI. An associated trust model IMPACTS (Intention, Measurability, Performance, Adaptivity, Communications, Transparency, and Security) will also be introduced to enable the practitioners to foster an assured and calibrated trust relationship between humans and their partner autonomous systems. A real-world example of human-autonomy teaming in a military context will be explained to illustrate the utility and effectiveness of these trust enablers.

Keywords: *trust, human factors, autonomous systems, interaction-centered design, human-autonomy teaming, human-machine symbiosis technology*

About the Keynote Speaker

Dr. Ming Hou received his PhD in Human Factors from the University of Toronto, Canada in 2002. He is currently a Senior Defence Scientist at Defence Research & Development Canada (DRDC) and the Principal Authority of Human-Technology Interactions within the Department of National Defence (DND), Canada where he received the prestigious DRDC Science and Technology Excellence Award in 2020. Dr. Hou is responsible for delivering technological solutions, science-based advice, and evidence-based policy recommendations to senior decision makers within DND and the Canadian Armed Forces (CAF) and their partner organizations. He also provides advice about the investment in and application of advanced technologies and methodologies for human-machine systems requirements and for AI and Autonomy science, technology and innovation strategies to the CAF and DND. He is an Integrator of the Canadian government $1.6B IDEaS program with responsibilities for guiding national R&D activities in AI, Automation, Robotics, and Telepresence. Dr. Hou is the Co-Chair of Human Factors Specialist Committee within NATO Joint Capability Group on Unmanned Aircraft Systems (UAS). His book: "Intelligent Adaptive Systems: An Interaction-Centered Design Perspective" provided guidance for the development of NATO STANRECs on "Human Systems Integration Guidance for UAS", "Sense and Avoid Guidance for UAS", and "UAS Human Factors Experimentation Guidebook". Dr. Hou also serves for multiple international scientific and technical associations/programs as a chair and a board member.

Perspectives on the Emerging Field of Autonomous Systems and its Theoretical Foundations

Yingxu Wang[1], *Fellow, IEEE,* Konstantinos N. Plataniotis[2], *Fellow, IEEE,* Arash Mohammadi[3], *SM, IEEE,* Lucio Marcenaro[4], *SM, IEEE,* Amir Asif[5], *SM, IEEE,* Ming Hou[6], *SM, IEEE,* Henry Leung[7], *Fellow, IEEE,* and Marina Gavrilova[8] *SM, IEEE*

[1,7] FIEEEs, Dept. of Electrical & Software Engineering
Schulich School of Engineering and Hotchkiss Brain Institute
Int'l Institute of Cognitive Informatics & Cognitive Computing (I2CICC)
University of Calgary, Canada
Emails: yingxu@ucalgary.ca and leungh@ucalgary.ca

[2] FIEEE, Dept. of Electrical & Computer Engineering
University of Toronto, ON, Canada
Email: kostas@ece.utoronto.ca

[3,5] SMIEEEs, Dept. of Computer Science
Concordia University, Montreal, Canada
Emails: arash.mohammadi@co ncordia.ca and amir.asif@concordia.ca

[4] SMIEEE, Dept. of DITEN, University of Genova, Italy
Email: lucio.marcenaro@unige.it

[6] SMIEEE, Toronto Research Centre, DRDC, Canada
Email: ming.hou@drdc-rddc.gc.ca

[8] SMIEEE, Dept. of Computer Science
University of Calgary, Canada
Emails: mgavrilo@ucalgary.ca

Abstract — Autonomous systems are advanced intelligent systems and general AI technologies triggered by the transdisciplinary development in intelligence science, system science, brain science, cognitive science, robotics, computational intelligence, and intelligent mathematics. AS are driven by the increasing demands in the modern industries of cognitive computers, deep machine learning, robotics, brain-inspired systems, self-driving cars, internet of things, and intelligent appliances. This paper presents a perspective on the framework of autonomous systems and their theoretical foundations. A wide range of application paradigms of autonomous systems are explored.

Keywords **—** Autonomous systems, intelligence science, system science, intelligent signal processing, general AI theory, brain-inspired systems

I. Introduction

Autonomous systems (AS) are an emerging field of advanced computational intelligence triggered by transdisciplinary developments in intelligence, cognitive, and system sciences as well as intelligent signal processing theories [1-12]. AS enable humans to involve in-the-loop of intelligent systems in order to coherently augment both human and machine intelligence to the maximum. AS lead to a General AI (GAI) theory [7, 20] for advancing machine intelligence.

AS refer to intelligent systems that "exhibit goal-oriented and potentially unpredictable and non-fully deterministic behaviors" by NATO [11]. In basic studies of intelligence science and systems science, the field of AS investigates intelligent systems for implementing advanced human intelligence by computational systems, neural networks, deep machine learning, and Intelligent Mathematics (IM) [14], which embodies high-level machine intelligence built on those of imperative and adaptive systems.

It is recognized that the theoretical foundations for AI in general, and for AS in particular, were not sufficiently mature in the past 60 years for intelligent engineering. As a consequence, few fully autonomous systems have been developed [5, 9, 10, 11]. The state-of-the-art of AI systems is still bounded by the intelligence bottleneck of adaptive systems where machine intelligence is constrained by the low-level reflexive, imperative, and deterministic intelligent abilities [7].

The transdisciplinary advances in intelligence, cognition, computer, signal/sensor, and system sciences have triggered the emerging field of AS [5, 6, 10]. The ultimate goal of AS is to implement a brain-inspired system that may think and act as a human counterpart in hybrid intelligent systems. AS are driven by the increasing demands in the modern industries of cognitive computer, deep machine learning, robotics, brain-inspired systems, self-driving cars, internet of things, and intelligent appliances [6].

This paper explores the nature and the theoretical framework of AS beyond traditional reflexive, imperative, and adaptive systems. A hierarchical intelligence model is introduced in Section II to elaborate the evolution of human and system intelligence as a recursive structure and an inductive process. The theoretical foundations of AS are formally described in Section III by a recursive mathematical model of AS. Then, the framework of IEEE ICAS'21 program is presented in Section IV which represents the co-chairs' perspectives on the inaugural conference series of AS and engineering applications.

II. The Emergence of Autonomous Systems

The transdisciplinary advances towards AS are explored in this section, which seeks what kinds of structural and behavioral

properties may constitute the intelligence power of AS beyond traditional systems. It explains how system intelligence aggregates from reflexive, imperative, adaptive intelligence to autonomous and cognitive intelligence.

2.1 From Reflexive, Imperative, Adaptive Systems to Autonomous and Cognitive Systems

Intelligence is a paramount cognitive ability of humans that may be mimicked by computational intelligence and AS. *Intelligence science* is a contemporary discipline that studies the mechanisms and properties of intelligence, and the theories of intelligence across the neural, cognitive, functional, and mathematical levels from the bottom up [3, 9, 14, 20]. Therefore, the level of intelligent is a key characteristic for distinguishing if a system is an AS underpinned by intelligence science.

Definition 1. *Intelligence* $\dot{\mathbb{I}}$ is a human, animal, or system ability that autonomously transfers a piece of information I into a behavior B (to-do) or an item of knowledge K (to-be), particularly the former:

$$\dot{\mathbb{I}} = \begin{array}{l} f_{to-do} : I \to B \\ \mid \; f_{to-be} : I \to K \end{array} \tag{1}$$

A classification of intelligent systems may be derived based on the forms of inputs and outputs dealt by the system as shown in Table 1. The reflexive or imperative systems are capable to process deterministic stimuli by deterministic or indeterministic algorithms, respectively. The adaptive systems are designed for dealing with indeterministic stimuli by deterministic behaviors predefined at design time. However, AS are characterized by both indeterministic stimuli and indeterministic (problem-specific or goal-oriented) behaviors pending for run-time contexts.

Table 1. Characteristics of autonomous and nonautonomous systems

Intelligent behaviors		**Behavior (O)**	
		Deterministic	Indeterministic
Stimulus (I)	Deterministic	*Reflexive system*	*Imperative system*
	Indeterministic	*Adaptive system*	*Autonomous system*

Definition 2. *Autonomous systems* (*AS*) are advanced intelligent systems that function without human intervention for implementing complex cognitive abilities aggregating from reflexive, imperative, and adaptive intelligence to autonomous and cognitive intelligence.

AS is an indeterministic nonlinear system that depends not only on current stimuli or demands, but also on internal status, willingness, and knowledge formed by long-term historical events and current rational or emotional goals. AS implements nondeterministic, context-dependent, and adaptive behaviors closer to the level of human cognitive intelligence.

2.2 The Hierarchical Model of Intelligence for AS

A *Hierarchical Intelligence Model* (HIM) is introduced to classify the levels of intelligence and their recursive properties in intelligence science as illustrated in Figure 1 based on the *abstract intelligence* (αI) theory [20]. As shown in Figure 1, the levels of natural and system intelligence may be aggregated from those of reflexive, imperative, adaptive, autonomous, and cognitive intelligence with 16 categories of intelligent behaviors. Types of system intelligence across the HIM layers are explained in the following subsections using the *event-dispatching mechanism* [18] as defined in Eq. (2). Rigorous mathematical models will be formally described in Section III.

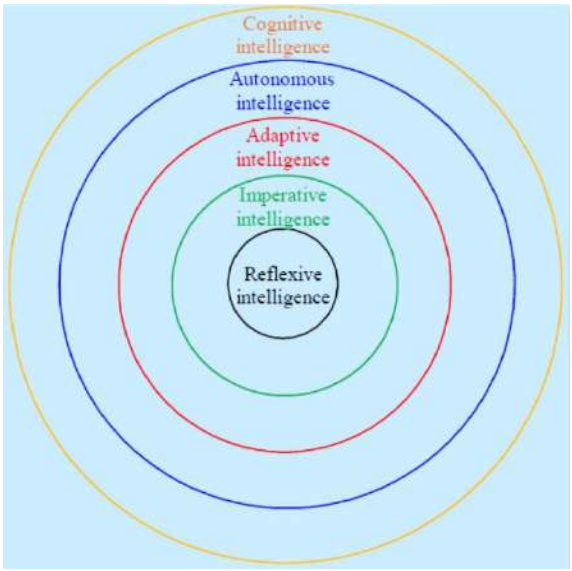


Fig. 1. The Hierarchical Intelligence Model (HIM)

1) *Reflexive intelligence* $\dot{\mathbb{I}}_{ref}$ is the bottom-layer intelligence of AS coupled by a stimulus and a reaction. $\dot{\mathbb{I}}_{ref}$ is shared among humans, animals, and machines, which forms the foundation of higher layer intelligence. $\dot{\mathbb{I}}_{ref}$ is a set of wired behaviors directly driven by specifically coupled external stimuli or trigger events.

2) *Imperative intelligence* $\dot{\mathbb{I}}_{imp}$ is a form of instructive and reflective behaviors dispatched by a system based on the layer of reflexive intelligence. $\dot{\mathbb{I}}_{imp}$ encompasses event-driven behaviors, time-driven behaviors, and interrupt-driven behaviors. The imperative system powered by $\dot{\mathbb{I}}_{imp}$ is not adaptive yet, and may merely implement deterministic, context-free, and stored-program controlled behaviors as a classical stored-program-controlled system.

3) *Adaptive intelligence* $\dot{\mathbb{I}}_{adp}$ is a form of run-time determined behaviors where a set of predictable scenarios is determined for processing variable problems. $\dot{\mathbb{I}}_{adp}$ encompasses analogy-based behaviors, feedback-modulated behaviors, and environment-awareness behaviors. $\dot{\mathbb{I}}_{ada}$ is constrained by deterministic rules where the scenarios are prespecified in design-time. If a request is out of the defined domain of an adaptive system, its behaviors will no longer be adaptive or predictable.

4) *Autonomous intelligence* $\dot{\mathbb{I}}_{aut}$ is the 4th-layer intelligence powered by internally motivated and self-generated behaviors underpinned by senses of system consciousness and environment awareness. $\dot{\mathbb{I}}_{aut}$ encompasses the perceptive behaviors, problem-driven behaviors, goal-oriented behaviors, decision-driven behaviors, and deductive behaviors built on Layers 1 through 3 intelligent behaviors. $\dot{\mathbb{I}}_{aut}$ is self-driven by the system based on internal consciousness and environmental awareness beyond the deterministic behaviors of adaptive intelligence. $\dot{\mathbb{I}}_{aut}$ represents nondeterministic, context-dependent, run-time autonomic, and self-adaptive behaviors.

5) *Cognitive intelligence* $\dot{\mathbb{I}}_{cog}$ is the 5th-layer of intelligence that generates inductive- and inference-based behaviors powered by autonomous reasoning. $\dot{\mathbb{I}}_{cog}$ encompasses the knowledge-based behaviors, learning-driven behaviors, inference-driven behaviors, and inductive behaviors built on the intelligence powers of Layers 1 through 4. $\dot{\mathbb{I}}_{cog}$ is nonlinear, nondeterministic, context-dependent, knowledge-dependent, and self-constitute, which represents the highest level of system intelligence mimicking the brain.

The mathematical models of HIM explain why the current level of machine intelligence had been stuck at the level of adaptive intelligence in the past 60 years, because of the lack of matured theories and mathematical means for implementing fully autonomous and cognitive intelligence comparable to human natural intelligence.

III. The Theoretical Foundations of Autonomous Systems

The framework of AS in Section II reveals that *autonomy* is a property of intelligent and cognitive systems that may change their behaviors in response to unanticipated events and unclear causality without human intervention driven by autonomous decision-making beyond predetermined adaptive behaviors. AS implements nondeterministic, context-dependent, and self-adaptive behaviors dependent not only on current stimuli or demands, but also on internal status and willingness formed by long-term historical events and current rational or emotional goals. The major capabilities of AS will need to be extended to the cognitive intelligence level towards highly intelligent systems beyond classic adaptive and imperative systems.

On the basis of the HIM model, a set of generic mathematical models of AS may be introduced as a rigorous theory towards designed and implementation of various AS paradigms.

Definition 3. The *mathematical model of AS* is a high-level intelligent system for implementing advanced intelligent abilities compatible to human intelligence in systems as an *event* ($e|\mathrm{S}$) - *behavior* ($B|\mathrm{PM}$) dispatching mechanism $@e|\mathrm{S} \hookrightarrow B|\mathrm{PM}$:

$$AS \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{AS}} @e_{AS}^{i}|\mathrm{S} \hookrightarrow [B_{AS}(i)|\mathrm{PM} \mid B_{AS}(i)|\mathrm{PM} \geq 4] \quad (2)$$

which extends the power of system intelligence from reflexive, imperative, and adaptive to autonomous and cognitive intelligence, where the *big-R* calculus denotes recurrent structures or iterative behaviors [22].

The Mathematical Model of Hierarchical AS Intelligence

LevelsOfIntelligence|SM $\triangleq$

{**1. Reflexive intelligence** (wired behaviors)

1.1 $\dot{\mathbb{I}}_{ref} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{ref}} @e_i|\mathrm{REF} \hookrightarrow B_{ref}(i)|\mathrm{PM}$ // Sensory-driven intelligence

2. Imperative intelligence (predefined event-driven behaviors)

2.1 $\dot{\mathbb{I}}_{imp}^{e} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_e} @e_i|\mathrm{E} \hookrightarrow B_{imp}^{e}(i)|\mathrm{PM}$ // Event-driven intelligence

2.2 $\dot{\mathbb{I}}_{imp}^{t} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_t} @e_i|\mathrm{TM} \hookrightarrow B_{imp}^{t}(i)|\mathrm{PM}$ // Time-driven intelligence

2.3 $\dot{\mathbb{I}}_{imp}^{int} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{int}} @e_i|\mathrm{INT} \hookrightarrow B_{imp}^{int}(i)|\mathrm{PM}$ // Interrupt-driven intelligence

3. Adaptive intelligence (run-time determined behaviors)

3.1 $\dot{\mathbb{I}}_{adp}^{ab} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{ab}} @e_i|\mathrm{AR} \hookrightarrow B_{adp}^{ab}(i)|\mathrm{PM}$ // Analogy-based intelligence

3.2 $\dot{\mathbb{I}}_{adp}^{fm} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{fm}} @e_i|\mathrm{FM} \hookrightarrow B_{adp}^{fm}(i)|\mathrm{PM}$ // Feedback-modulated intel.

3.3 $\dot{\mathbb{I}}_{adp}^{ea} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{ea}} @e_i|\mathrm{EA} \hookrightarrow B_{adp}^{ea}(i)|\mathrm{PM}$ // Environment-aware intel.

4. Autonomous intelligence (self-driven behaviors)

4.1 $\dot{\mathbb{I}}_{aut}^{pe} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{pe}} @e_i|\mathrm{PE} \hookrightarrow B_{aut}^{pe}(i)|\mathrm{PM}$ // Perceptive intelligence

4.2 $\dot{\mathbb{I}}_{aut}^{pd} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{pd}} @e_i|\mathrm{PD} \hookrightarrow B_{aut}^{pd}(i)|\mathrm{PM}$ // Problem-driven intelligence

4.3 $\dot{\mathbb{I}}_{aut}^{go} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{go}} @e_i|\mathrm{GO} \hookrightarrow B_{aut}^{go}(i)|\mathrm{PM}$ // Goal-driven intelligence

4.4 $\dot{\mathbb{I}}_{aut}^{dd} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{dd}} @e_i|\mathrm{DD} \hookrightarrow B_{aut}^{dd}(i)|\mathrm{PM}$ // Decision-driven intel.

4.5 $\dot{\mathbb{I}}_{aut}^{de} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{de}} @e_i|\mathrm{DE} \hookrightarrow B_{aut}^{de}(i)|\mathrm{PM}$ // Deductive intelligence

5. Cognitive intelligence (learning and inference-based behaviors)

5.1 $\dot{\mathbb{I}}_{cog}^{kb} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{kb}} @e_i|\mathrm{KB} \hookrightarrow B_{cog}^{kb}(i)|\mathrm{PM}$ // Knowledge-based intel.

5.2 $\dot{\mathbb{I}}_{cog}^{ld} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{ld}} @e_i|\mathrm{LD} \hookrightarrow B_{cog}^{ld}(i)|\mathrm{PM}$ // Learning-driven intel.

5.3 $\dot{\mathbb{I}}_{cog}^{if} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{if}} @e_i|\mathrm{IF} \hookrightarrow B_{cog}^{if}(i)|\mathrm{PM}$ // Inference-driven intel.

5.4 $\dot{\mathbb{I}}_{cog}^{id} \triangleq \mathop{\mathbb{R}}_{i=1}^{n_{id}} @e_i|\mathrm{ID} \hookrightarrow B_{cog}^{id}(i)|\mathrm{PM}$ // Inductive intelligence

}

Fig. 2. The mathematical framework of hierarchical AS intelligence

According to the HIM model, the *behavioral model* of AS is inclusively aggregated from the bottom up among $AS|\S \triangleq (B_{Ref}, B_{Imp}, B_{Adp}, B_{Aut}, B_{Cog})$, where |§ denotes a system embodied by the set of reflexive, imperative, adaptive, autonomous, and cognitive behaviors as shown in Figure 2.

Theorem 1. An *AS* is characterized by a) a *recursively hierarchical* architecture and b) a series of *recursively inclusive* behaviors:

$$AS|\S \triangleq \begin{cases} a)\ \underset{k=1}{\overset{4}{R}} B^k(B^{k-1}), B^0 = \underset{i=1}{\overset{n_{ref}}{R}} @e_i|\text{REF} \hookrightarrow B_{ref}(i)|\text{PM} \\ b)\ B_{Cog} \supset B_{Aut} \supset B_{Ada} \supset B_{Imp} \supset B_{Ref} \end{cases} \quad (3)$$

Proof. $\forall AS|\S$, a) The recursive behavioral architecture $\underset{k=1}{\overset{4}{R}} B^k(B^{k-1})$ is necessary to aggregate the *AS*' functions through B^0 to B^4 from the bottom up, *iff* B^0 is deterministic; b) Because the five-level behaviors are in a partial order, the *recursive inclusivity* across all layers of behaviors is sufficient for composing the AS. ■

Theorem 1 indicates that any lower layer behavior of AS is a subset of those of a higher layer. In other words, any higher layer behavior of AS is a natural aggregation of those of lower layers as shown in Figure 2. According to the necessary and sufficient conditions stated in Theorem 1, a hybrid AS with humans in the loop will gain strengths towards the implementation of cognitive intelligent systems. The cognitive AS will sufficiently enable a powerful GAI system with the strengths of both human and machine intelligence. This is what intelligence and system sciences may inspire towards the development of fully autonomous systems in highly demanded engineering applications.

The HIM model and Theorem 1 reveal the ultimate goal of AI and machine intelligence. They lead to the finding of the 6th and most important form of machine learning known as *cognitive knowledge learning* [17] beyond traditional learning technologies for object identification, cluster classification, pattern recognition, functional regression and behavior generation (gaming) [19]. They also enabled the discovery that the basic unit of knowledge is a *binary relation* (*bir*) [17].

IV. The Framework of IEEE ICAS'21

The framework of the inaugural IEEE International Conference on Autonomous Systems (ICAS'21) is highlighted in Table 2 where three themes are covered in the categories of theoretical foundations of AS, emerging fields of AS, and engineering paradigms.

Advances in *AS* are expected to pave a way towards highly intelligent machines for augmenting human capabilities. Typical emerging *AS* include unsupervised computational intelligence, cognitive systems, brain-inspired systems, general automobiles, unmanned systems, human intelligence augmentation systems, intelligent defence systems, and intelligent IoTs.

Table 2. The Program Framework of IEEE ICAS'21

Theoretical Foundations of AS	Emerging Fields of AS	AS Engineering
• Intelligent foundations of AS	• Autonomous computers	• Applied paradigms of AS
• System foundations of AS	• Autonomous algorithms	• Autonomous programming
• Mathematical foundations of AS	• Brain-inspired AS	• Cognitive inference Engines
• Computational foundations of AS	• Autonomous machine learning	• Autonomous robots
• Brain science foundations of AS	• Autonomous IoTs	• Distributed AS
• Cognitive foundation of AS	• Self-driving vehicles and vessels	• Embedded AS
• Bottlenecks of adaptive Systems	• Autonomous robots	• Communications among AS
• Indeterministic and uncertainty behaviors of AS	• Real-time AS	• Communications Between AS and humans
• Interaction between humans and AS	• Autonomous unmanned systems	• Autonomous operating Systems
• Autonomous Computing platforms	• Trustworthiness of AS	• Autonomous sensors
• Neurological foundations of AS	• Mission critical systems	• Autonomous swarms
• Signal processing theories of AS	• Autonomous perception/awareness	• Social AS

None of the *AS* applications is trivial towards the next generation of cognitive computers, *GAI*, and hybrid symbiotic human-machine societies. Recent *AS* projects undertaken in our labs address challenges for abstract intelligence, intelligent mathematics for *AS*, the tripartite framework of *AS* trustworthiness, autonomous decision making, a transdisciplinary theory for cognitive cybernetics, humanity, and systems science, cognitive foundations of knowledge science, and the abstract intelligence theory for *AS* [1, 20-26]. The advances of *AS* theories and technologies will lead to the era of intelligence revolution for unprecedented breakthroughs to enable pervasive AS, which help to augment human intelligent power by autonomous and cognitive intelligence.

V. Conclusion

It has been recognized that autonomous systems are emerged from perceptive, problem-driven, goal-driven, decision-driven, and deductive intelligence. This work has explored basic research on the intelligence and system foundations of autonomous systems. A Hierarchical Intelligence Model (HIM) has been developed for elaborating the properties of autonomous systems built upon reflexive, imperative, and adaptive systems. The nature of system autonomy and human in-the-loop of autonomous systems has been formally analyzed. This work has provided a theoretical framework for developing cognitive autonomous systems towards highly demanded engineering applications including brain-inspired cognitive systems, unmanned systems, self-driving vehicles, cognitive robots, and intelligent IoTs.

ACKNOWLEDGEMENT

This work is supported in part by the DND IDEaS AutoDefence project, NSERC of Canada, and the IEEE SPS Autonomous System Initiative (ASI). The authors would like to thank the anonymous reviewers for their valuable suggestions and comments on this paper.

REFERENCES

[1] V. Mnih et al. (2015), Human-level Control through Deep Reinforcement Learning. *Nature*, 518: 529–533.

[2] E.A. Bender (2000), *Mathematical Methods in Artificial Intelligence*, IEEE CS Press, Los Alamitos, CA.

[3] G.J. Klir (1992), *Facets of Systems Science*, Plenum, NY.

[4] T. O'connor and D. Robb *eds.* (2003), *Philosophy of Mind: Contemporary Readings*, Routledge, London, UK.

[5] Y. Wang, M. Hou, K.N. Plataniotis, S. Kwong, H. Leung, E. Tunstel, I.J. Rudas, and L. Trajkovic (2021), Towards a Theoretical Framework of Autonomous Systems Underpinned by Intelligence and Systems Sciences, *IEEE/CAS Journal of Automatica Sinica*, 8(1), 52-63.

[6] K. Grise, T. Martinez, and R. Saracco (2021), The Winding Path towards Symbiotic Autonomous Systems, *Philosophical Transactions of Royal Society (A)*, Oxford, UK, in press.

[7] Y. Wang, F. Karray, O. Kaynak, S. Kwong, H. Leung, K.N. Plataniotis, M. Hou, I.J. Rudas, E. Tunstel, L. Trajkovic, and J. Kacprzyk (2021), Perspectives on the Philosophical, Cognitive and Mathematical Foundations of Symbiotic Autonomous Systems (SAS), *Philosophical Transactions of Royal Society (A)*, Oxford, UK, 379(2207):1-20.

[8] R.A. Wilson and C.K. Frank eds. (2001), *The MIT Encyclopedia of the Cognitive Sciences*, MIT Press, MA.

[9] D.P. Watson and D.H. Scheidt (2005), Autonomous Systems, *Johns Hopkings Appl. Tech. Digest*, 26(4), pp. 268-376.

[10] Y. Wang, S. Kwong, H. Leung, J. Lu, M.H. Smith, L. Trajkovic, E. Tunstel, K.N. Plataniotis, G. Yen, and W. Kinsner (2019), Brain-Inspired Systems: A Transdisciplinary Exploration on Cognitive Cybernetics, Humanity, and Systems Science towards AI, *IEEE System, Man and Cybernetics Magazine*, 5(3): 6-13.

[11] M. Hou, S. Banbury and C. Burns (2014), *Intelligent Adaptive Systems: An Interaction-Centered Design Perspective*, CRC Press, NY.

[12] A. Leeper, K. Hsiao, M. Ciocarlie, L. Takayama, D. Gossow (2012), Strategies for Human-in-the-Loop Robotic Grasping, *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 1-8.

[13] J. Albus, J. (1991), Outline for a Theory of Intelligence, *IEEE Transactions on Systems, Man and Cybernetics*, 21(3), 473- 509.

[14] Y. Wang (2020), Keynote: Intelligent Mathematics: A Basic Research on Foundations of Autonomous Systems, General AI, Machine Learning, and Intelligence Science, *IEEE 19th Int'l Conf. on Cognitive Informatics and Cognitive Computing* (ICCI*CC'20), Tsinghua Univ., Beijing, China, Sept., p.5.

[15] Y. Wang (2003), On Cognitive Informatics, *Brain and Mind: A Transdisciplinary Journal of Neuroscience and Neurophilosophy*, 4(2), 151-167.

[16] Mohammadi, A. and K.N. Plataniotis (2017), Event-Based Estimation with Information-Based Triggering and Adaptive Update, *IEEE Transactions on Signal Processing*, 65(18), pp. 4924-4939.

[17] Y. Wang (2017), Keynote: Cognitive Foundations of Knowledge Science and Deep Knowledge Learning by Cognitive Robots, 16th IEEE International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC 2017), University of Oxford, UK, IEEE CS Press, July, p. 4.

[18] Y. Wang (2010), Cognitive Robots: A Reference Model towards Intelligent Authentication, *IEEE Robotics and Automation*, 17(4), pp. 54-62.

[19] Y. LeCun, Y., Y. Bengio and G.E. Hinton (2015), Deep Learning, *Nature*, 521(7553):436-444.

[20] Y. Wang (2009), On Abstract Intelligence: Toward a Unified Theory of Natural, Artificial, Machinable, and Computational Intelligence, *International Journal of Software Science and Computational Intelligence*, Jan., 1(1): 1-17.

[21] A. Poursaberi, S. Yanushkevich, M. Gavrilova, et al. (2013), Situational awareness through biometrics, *IEEE Computer*, 46(5), pp. 102–104.

[22] Y. Wang (2008), On the Big-R Notation for Describing Interactive and Recursive Behaviors, *International Journal of Cognitive Informatics and Natural Intelligence*, 2(1):17-28.

[23] Y. Wang (2015), Concept Algebra: A Denotational Mathematics for Formal Knowledge Representation and Cognitive Robot Learning, *Journal of Advanced Mathematics and Applications*, 4(1):61-86.

[24] Y. Wang (2009), Formal Description of the Cognitive Process of Memorization, *Transactions on Computational Science*, v. 5, Springer, pp. 81-98.

[25] Y. Wang (2012), On Visual Semantic Algebra (VSA): A Denotational Mathematical Structure for Modeling and Manipulating Visual Objects and Patterns, *Software and Intelligent Sciences: New Transdisciplinary Findings*, pp.68-81.

[26] Y. Wang, D. Liu and G. Ruhe (2004), Formal Description of the Cognitive Process of Decision Making, *Proceedings of the 3rd IEEE International Conference on Cognitive Informatics*, IEEE CS, Press, pp. 124-130.

OPTIMAL MULTIDIMENSIONAL CYCLIC CONVOLUTION ALGORITHMS FOR DEEP LEARNING AND COMPUTER VISION APPLICATIONS

Prof. Ioannis Pitas

pitas@csd.auth.gr
Department of Informatics, Aristotle University of Thessaloniki
Thessaloniki 54124, Greece

ABSTRACT

1D, 2D and multidimensional convolutions are basic tools in deep learning, notably in convolutional neural networks (CNNs) and in computer vision (template matching, correlation trackers). Therefore, fast 1D/2D/3D convolution algorithms are essential for advanced machine learning and computer vision. This paper presents: 1) novel optimal n-D cyclic convolution algorithms having minimal multiplicative complexity that are much faster than any competing convolution algorithm internationally and 2) methods for speeding up such optimal convolution algorithms on GPUs and multicore CPUs. Such a speedup is very important both for CNN training and CNN testing, particularly in embedded environments (e.g., on drones) and real-time applications (e.g., fast CNN inference for object detection and correlation trackers for embedded real-time object tracking).

Index Terms— Convolutional Neural Networks, Fast convolutions

1. INTRODUCTION

2D convolutional layers in CNNs [1] typically convolve the feature map $\mathbf{X}_l$ (3D tensor of dimensions $N_l \times M_l \times C_l$) of the neural network layer l with a k-th convolution kernel $\mathbf{w}_{l,k}$ (3D tensor of dimensions $H_{l,k} \times W_{l,k} \times C_l$), add a bias term $b(l,k)$ and then pass it through a nonlinearity activation function f, (e.g., RELU), to produce the feature map $\mathbf{X}_{l+1,k}$ (3D tensor of dimensions $N_{l+1,k} \times M_{l+1,k} \times C_{l+1}$) of layer $l+1$:

$$x(i,j,c_{l+1},l+1,k) = f(b(l,k)+ \sum_{c=1}^{C_l} \sum_{i'=0}^{H_{l,k}} \sum_{j'=0}^{W_{l,k}} h(i',j',l,k)x(i-i',j-j',c,l,k)). \quad (1)$$

The extension to higher spatial or spatiotemporal dimensions is straightforward. It is essential to devise fast 2D and multidimensional convolution algorithms, in order to have fast CNN training and testing. Furthermore, the same algorithms can be used for calculating the 2D correlation of an input image x) and a template h, e.g., for fast correlation trackers [2], [3].

The construction of fast convolution algorithms is a heavily researched topic in the signal processing community. It reached maturity in the 90ties [4], [5], [6].

Recently, a resurgence of fast linear convolution algorithms for CNNs occured, collectively called Winograd convolutions [7], [8]. Various implementations for GPUs and multicore CPUs appeared [7], [9] and numerical stability issues have been investigated. However, most of the recent algorithms are suboptimal.

2. FAST 2D AND MULTIDIMENSIONAL CONVOLUTION ALGORITHMS WITH MINIMAL COMPUTATIONAL COMPLEXITY

A linear convolution of signal x having length L with a convolutional kernel h having length M produces an output signal $y(n) = x(n) * h(n)$ of length $L+M-1$ and can be embedded in a cyclic convolution of length $N \geq L+M-1$, by zero padding both the signal x and convolution kernel h. The following relation holds for the Z-transform of an 1D cyclic convolution of length N:

$$y(n) = x(n) \circledast h(n) <=> Y(z) = X(z)H(z) \mod (Z^N - 1). \quad (2)$$

The 1D cyclic Winograd convolution algorithms are proven to be of the form:

$$\mathbf{y} = \mathbf{C}(\mathbf{Ax} \otimes \mathbf{Bh}). \quad (3)$$

This fast convolution architecture is shown in Figure 1.

In this section, we shall focus on fast 2D cyclic convolution:

$$y(k_1,k_2) = \sum_{i_1}^{N_1} \sum_{i_2}^{N_2} h(i_1,i_2)x((k_1-i_1)_{N_1},(k_2-i_2)_{N_2}) \quad (4)$$

algorithms having minimal computational complexity, as the methodology for fast multidimensional convolutions is similar [5], [6]. These convolutions have the following form in

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871479 (AERIAL-CORE).

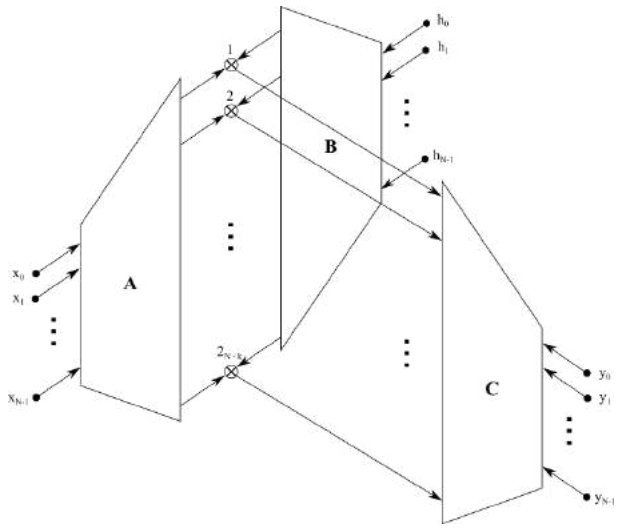


Fig. 1: 1D Winograd cyclic convolution algorithm.

the 4D Z domain (z_1, z_2):

$$Y(z_1, z_2) = X(z_1, z_2)H(z_1, z_2) \mod ({z_1}^{N_1} - 1), ({z_2}^{N_2} - 1), \tag{5}$$

or more generally:

$$Y(z_1, z_2) = X(z_1, z_2)H(z_1, z_2) \mod P_1(z_1), P_2(z_2) \tag{6}$$

If $P_i(z_i), i = 1, 2$ factorize as follows:

$$P_i(z_i) = \prod_{j_i=1}^{\nu_i} P_{ij_i}(z_i), \quad 1 \le j_i \le \nu_i, \quad i = 1, 2, \tag{7}$$

the convolution (6) is split into $\nu_1\nu_2$ smaller products:

$$Y_{1j_1,2j_2}(z_1, z_2) = X_{1j_1,2j_2}(z_1, z_2)H_{1j_1,2j_2}(z_1, z_2) \tag{8}$$

$$\mod P_{1j_1}(z_1), P_{2j_2}(z_2), 1 \le j_i \le \nu_i, i = 1, 2. \tag{9}$$

By finding the polynomials:

$$R_{ij_i}(z_i) = \delta_{j_i,k_i} \mod P_{ij_i}(z_i), P_{ik_i}(z_i) \quad 1 \le j_i \le \nu_i, k_i \le \nu_i, \quad i = 1, 2, \tag{10}$$

we can reconstruct $Y(z_1, z_2)$ as follows [6]:

$$Y(z_1, z_2) = \sum_{j_1=1}^{\nu_1} Y_{1j_1,2j_2}(z_1, z_2) \sum_{j_2=1}^{\nu_2} R_{1j_1}(z_1)R_{2j_2}(z_2) \mod P_1(z_1), P_2(z_2). \tag{11}$$

The algorithm (6)-(11) is essentially the split nesting convolution algorithm [4], which is a variation of the nesting algorithm of [10].

The computational complexity of this 2D convolution algorithm is $(2N_1 - \nu_1)(2N_2 - \nu_2)$, which is $O(N^2)$, i.e., much lower than the computational complexity of $O(N^4)$ of the 2D cyclic convolution computation by its definition (4). However, this is not the minimal computational complexity algorithm, as: a) each polynomial $P_{1j_1}(z_1), 1 \le j_1 \le \nu_1$ can possibly be further factorized in k_{j_1,j_2} factors over the field $Q[z_2]/P_{2j_2}(z_2), 1 \le j_2 \le \nu_2$ or b) vice versa, each polynomials $P_{2j_2}(z_2), 1 \le j_2 \le \nu_2$ can possibly be further factorized k_{j_2,j_1} factors over the fields $Q[z_1]/P_{1j_1}(z_1), 1 \le j_1 \le \nu_1$. By examining both these further factorizations (a, b) for each product (8), we can derive 2D cyclic convolution algorithms having minimal computational complexity [6]:

$$M = \sum_{j_1=1}^{\nu_1} \sum_{j_2=1}^{\nu_2} min((2N_{j_2} - 1)(2N_{j_1} - k_{j_1,j_2}), (2N_{j_1} - 1)(2N_{j_2} - k_{j_2,j_1})) \tag{12}$$

Such algorithms take the general form of (3). However, the construction of matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ requires good Algebra skills and is far from trivial. We constructed such novel optimal algorithms for several $N \times N$ cases, notably for $p \times p$ (e.g., for $p = 3, 5, 7$) and for $2^l \times 2^l$ (e.g., for 4×4) cyclic convolutions. An illustrative example of this procedure for a 3×3 cyclic convolution can be found in the next section.

CNNs typically employ 2D linear convolutions having small convolution kernels (e.g., below 11×11 coefficients). They, in turn, can be embedded in rather small 2D $N \times N$ cyclic convolution, as input images are typically split into small blocks to enable block-based 2D $N \times N$ cyclic convolution calculations that are highly (and easily) parallelizable in GPUs, as can be seen in a subsequent section.

3. EXAMPLE: FAST 2D 3×3 CYCLIC CONVOLUTION ALGORITHM HAVING MINIMAL COMPUTATIONAL COMPLEXITY

A 2D 3×3 cyclic convolution is defined as follows:

$$Y(z_1, z_2) = H(z_1, z_2)X(z_1, z_2) \mod z_1^3 - 1, z_2^3 - 1, \tag{13}$$

where:

$$X(z_1, z_2) = x_{00} + x_{01}z_2 + x_{02}z_2^2 + x_{10}z_1 + x_{11}z_1z_2 + x_{12}z_1z_2^2 + x_{20}z_1^2 + x_{21}z_1^2z_2 + x_{22}z_1^2z_2^2 \tag{14}$$

$$H(z_1, z_2) = h_{00} + h_{01}z_2 + h_{02}z_2^2 + h_{10}z_1 + h_{11}z_1z_2 + h_{12}z_1z_2^2 + h_{20}z_1^2 + h_{21}z_1^2z_2 + h_{22}z_1^2z_2^2. \tag{15}$$

The polynomial $z^3 - 1$ is analyzed as follows:

$$z^2 - 1 = (z - 1)(z^2 + z + 1). \tag{16}$$

Therefore, the 2D 3×3 cyclic convolution is decomposed as follows:

$$X_1(z_1, z_2) = X(z_1, z_2) \mod (z_1 - 1), (z_2 - 1) \quad (17)$$

$$X_2(z_1, z_2) = X(z_1, z_2) \mod (z_1 - 1)(z_2^2 + z_2 + 1) \quad (18)$$

$$X_3(z_1, z_2) = X(z_1, z_2) \mod (z_2 - 1)(z_1^2 + z_1 + 1) \quad (19)$$

$$X_4(z_1, z_2) = X(z_1, z_2) \mod (z_1^2 + z_1 + 1)(z_2^2 + z_2 + 1). \quad (20)$$

We notice that the polynomial $z_1^2 + z_1 + 1$ can be factorized in the field $Q[z_2]/z_2^2 + z_2 + 1$ as follows:

$$z_1^2 + z_1 + 1 = (z_1 - z_2)(z_1 + 1 + z_2). \quad (21)$$

Thus, $Y_4(z_1, z_2)$ can be further decomposed in two terms:

$$X_{4_1}(z_1, z_2) = X_4(z_1, z_2) \mod (z_1 - z_2)(z_2^2 + z_2 + 1) \quad (22)$$

$$X_{4_2}(z_1, z_2) = X_4(z_1, z_2) \mod (z_1 + z_2 + 1)(z_2^2 + z_2 + 1) \quad (23)$$

By employing CRT, $Y_4(z_1, z_2)$ is reconstructed from $Y_{4_1}(z_1, z_2)$ and $Y_{4_2}(z_1, z_2)$ as follows:

$$Y_4(z_1, z_2) = \sum_{n=1}^{2} R_i(z_1, z_2) Y_{4_i}(z_1, z_2) \mod (z_1^2 + z_1 + 1)(z_2^2 + z_2 + 1) \quad (24)$$

$$R_1(z_1, z_2) = -\frac{1}{3}[2z_1 z_2 + z_2 + 1] \quad (25)$$

$$R_2(z_1, z_2) = \frac{1}{3}[(2z_2 + 1)z_1 + z_2 + 2]. \quad (26)$$

Then $Y(z_1, z_2)$ is reconstructed as follows:

$$Y(z_1, z_2) = \sum_{i=1}^{4} R_i(z_1, z_2) Y_i(z_1, z_2) \mod (z_1^3 - 1)(z_2^3 - 1), \quad (27)$$

where:

$$R_1(z_1, z_2) = \frac{1}{9}(z_1^2 + z_1 + 1)(z_2^2 + z_2 + 1) \quad (28)$$

$$R_2(z_1, z_2) = -\frac{1}{9}(z_1^2 + z_1 + 1)(z_2^2 + z_2 - 2) \quad (29)$$

$$R_3(z_1, z_2) = -\frac{1}{9}(z_1^2 + z_1 - 2)(z_2^2 + z_2 + 1) \quad (30)$$

$$R_4(z_1, z_2) = \frac{1}{9}(z_1^2 + z_1 - 2)(z_2^2 + z_2 - 2). \quad (31)$$

This leads to fast 2D 3×3 cyclic convolution algorithm of the form:

$$\mathbf{y} = \mathbf{C}(\mathbf{A}\mathbf{x} \otimes \mathbf{B}\mathbf{h}), \quad (32)$$

where:

$$\mathbf{A} = \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 & 0 & 1 & 0 & 1 & -1 \\ 1 & 0 & -1 & -1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & -1 & 0 & 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 1 & -1 & -1 & 0 & 1 \end{bmatrix}, \quad (33)$$

This fast algorithm requires only 13 multiplications, while the computation of the 3×3 cyclic convolution using its definition (13) requires 81 multiplications. Also note that the two matrix-vector products and the point-wise vector product are highly parallelizable. Furthermore, as expected, matrix $\mathbf{A}$, $\mathbf{B}$ entries are $0, \pm 1$. Therefore, if we use the form (3), we have no 'multiplications in the matrix-vector products. Furthermore, additions/subtractions used in the matrix vector products, e.g., $\mathbf{X} = \mathbf{A}\mathbf{x}$ can be grouped in subsums that can reused. In the case of the $\mathbf{A}\mathbf{x}$ computation, the additions can be reduced from 68 to 40, as can be seen in the flow diagram of Figure 2. The constructions of algorithms of the form (3), taking all these optimizations into account, is novel, does pay off and it is far from trivial.

Transformation of each one of the 13 rows of matrix $\mathbf{A}$ into 3×3 submatrices leads to interesting visualization patterns, as can be seen in Figure 3. Essentially, each matrix row (but for the first one) produces an input 2D signal (image) $\mathbf{x}$ transformation on certain directions and frequency bands.

4. SPEEDING UP AND PARALLELIZATION OF CONVOLUTION ALGORITHMS

In case $L >> M$, the signal x can be split in blocks of length L_B. Then the linear convolution can be split in smaller length $N \geq L_B + M - 1$ convolutions using overlap-add or overlap-save methods [11]. This approach leads to very easily parallelizable convolution algorithms.

The proposed 2D convolution method was implemented on GPU cards for a 3×3 convolution kernel on 512×512 pixel input images, using an optimal 4×4 cyclic convolution algorithm combined with an overlap-save block-based approach employing 65536 2×2 image blocks (tiles). Its execution time was a mere 0.0809 ms. It is 4,77 times faster than the fastest cuDNN convolution (GEMM-0) and 11,33 times faster than the corresponding cuDNN Winograd linear convolution rou-

$$\mathbf{C} = \frac{1}{27}\begin{bmatrix} 3 & 3 & -6 & 3 & 3 & -6 & 3 & 3 & -6 & 3 & 3 & -6 & 3 \\ 3 & 3 & 3 & -6 & 3 & -6 & 3 & 3 & 3 & -6 & 3 & 3 & -6 \\ 3 & -6 & 3 & 3 & 3 & -6 & 3 & -6 & 3 & 3 & -6 & 3 & 3 \\ 3 & 3 & -6 & 3 & 3 & 3 & -6 & 3 & 3 & -6 & -6 & 3 & 3 \\ 3 & 3 & 3 & -6 & 3 & 3 & -6 & -6 & 3 & 3 & 3 & -6 & 3 \\ 3 & -6 & 3 & 3 & 3 & 3 & -6 & 3 & -6 & 3 & 3 & 3 & -6 \\ 3 & 3 & -6 & 3 & -6 & 3 & 3 & -6 & 3 & 3 & 3 & 3 & -6 \\ 3 & 3 & 3 & -6 & -6 & 3 & 3 & 3 & -6 & 3 & -6 & 3 & 3 \\ 3 & -6 & 3 & 3 & -6 & 3 & 3 & 3 & 3 & -6 & 3 & -6 & 3 \end{bmatrix}. \quad (34)$$

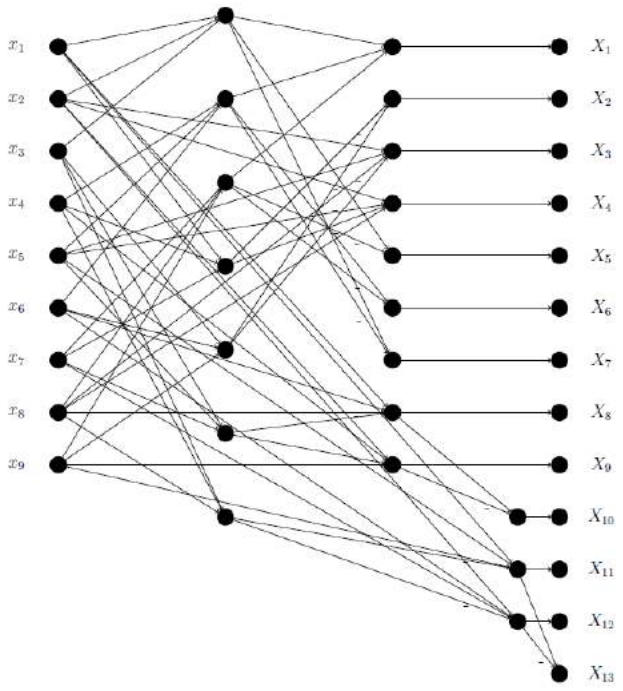


Fig. 2: Diagram for $\mathbf{Ax}$ calculation for a 3×3 Winograd cyclic convolution.

tine (Winograd-6) for the same convolution kernel size and image size. Therefore, the proposed algorithm is almost 5 times faster than the fastest known competing convolution algorithm internationally.

5. CONCLUSIONS

CNN are still slow to be used, e.g., for object detection in embedded vision systems [12]. Furthermore, R & D on CNNs moves towards higher dimensions for 3D spatial or spatiotemporal (video) analysis, where processing requirements are excessive. Finally, fast convolution structures and essential for fast object tracking (e.g., correlation trackers) in embedded systems [3]. Therefore, it indeed pays off to derive fast (possibly optimal) multidimensional convolution algorithms.

This paper presents novel 2D and n-D cyclic convolution algorithms having minimal computational complexity. They can be easily described by simple linear algebra operations with matrices having trivial entries $0, \pm 1$. Their structure is easily paralellizable. This renders these algorithms very attractive for multicore CPU and GPU processing. They are much faster than any competing convolution algorithm known today. However, they do have their drawbacks. In the general case of 2D $N \times N$ convolution, their structure cannot easily be obtained and requires strong mathematical skills. The process of automatically generating the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ to be used in (3) is an active and very interesting research topic.

6. REFERENCES

[1] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436, 2015.

[2] Joao F Henriques, Rui Caseiro, Pedro Martins, and Jorge Batista, "High-speed tracking with kernelized correlation filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 583–596, 2015.

[3] Paraskevi Nousi, Danai Triantafyllidou, Anastasios Tefas, and Ioannis Pitas, "Joint lightweight object tracking and detection for unmanned vehicles," in *Proceed-*

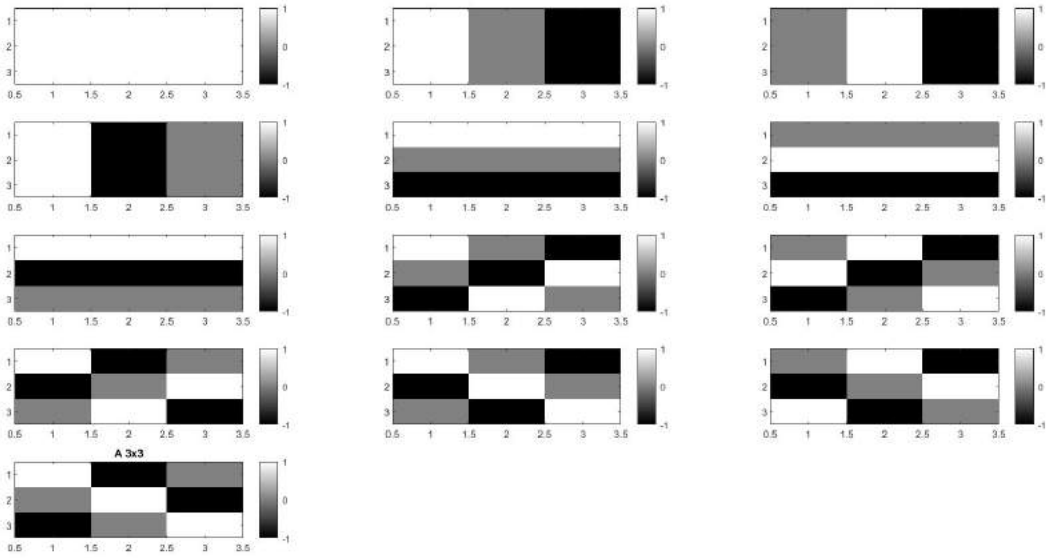


Fig. 3: Visualization of rows of the 3×3 Winograd cyclic convolution matrix $\mathbf{A}$.

ings of the of the IEEE International Conference on Image Processing (ICIP), 2019.

[4] Henri J Nussbaumer, *Fast Fourier transform and convolution algorithms*, Springer Science & Business Media, 1981.

[5] I Pitas, *Computational complexity study of multidimensional digital signal processing algorithms*, Aristotle University of Thessaloniki, Greece, 1985.

[6] Ioannis Pitas and M Strintzis, "Multidimensional cyclic convolution algorithms with minimal multiplicative complexity," *IEEE transactions on acoustics, speech, and signal processing*, vol. 35, no. 3, pp. 384–390, 1987.

[7] Sharan Chetlur, Cliff Woolley, Philippe Vandermersch, Jonathan Cohen, John Tran, Bryan Catanzaro, and Evan Shelhamer, "cudnn: Efficient primitives for deep learning," *arXiv preprint arXiv:1410.0759*, 2014.

[8] Andrew Lavin and Scott Gray, "Fast algorithms for convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4013–4021.

[9] Zhen Jia, Aleksandar Zlateski, Fredo Durand, and Kai Li, "Optimizing n-dimensional, winograd-based convolution for manycore cpus," in *Proceedings of the 23rd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. ACM, 2018, pp. 109–123.

[10] R Agarwal and J Cooley, "New algorithms for digital convolution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 5, pp. 392–410, 1977.

[11] Ioannis Pitas, *Digital image processing algorithms and applications*, John Wiley & Sons, 2000.

[12] N.Nikolaidis I. Mademlis, V. Mygdalis and I.Pitas, "Challenges in autonomous uav cinematography: An overview," *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 392–410, 2018.

INTERPRETABLE ANOMALY DETECTION USING A GENERALIZED MARKOV JUMP PARTICLE FILTER

Giulia Slavic, Pablo Marin, David Martin, Lucio Marcenaro, Carlo Regazzoni

University of Genova (Italy), University Carlos III de Madrid (Spain)

ABSTRACT

When performing anomaly detection on an autonomous vehicle's sensory data, it is fundamental to infer the cause of the found anomalies. This paper proposes a method for learning prediction models and detecting anomalies by decomposing the evolution of an agent's state into its different motion-related parameters. A filter is introduced based on Generalized Filtering to increase the interpretability of the results with respect to previous methods. The proposed anomaly detection method is tested on data from a real vehicle. We also consider the case in which multiple models are learned, how to extract the salient discriminatory features of each, and use the proposed anomaly detection method to perform behavior classification.

Index Terms— Anomaly detection, Kalman Filter, Particle Filter, Interpretable Machine Learning

1. INTRODUCTION

The learning of a model describing how the state of an agent evolves across time has many purposes: using the model to perform short-time or long-time prediction; classifying an agent based on what model it follows; performing anomaly detection distinguishing when the rules of the model are broken or when they are respected. Such applications are of interest in a variety of fields, from autonomous driving [1], to self-aware radios [2], to video surveillance [3], to weather forecasting [4], to medical image analysis [5].

When considering anomaly detection applied in particular to the self-aware agents' field, another important concept can be introduced, i.e., *interpretability*. It can be observed that there is no strict, universally recognized definition of interpretability, which is often also associated with explainability. In [6], interpretability is defined as answering the question "How does the model work?", and explainability as answering the question "What else can the model tell me?". It is interesting to note the desiderata and properties of interpretable research defined by [6], including causality, decomposability of the individual parameters, and algorithmic transparency. Among the interpretable algorithms also fall Bayesian models such as Dynamic Bayesian Networks (DBNs) [7,8], which are used in our paper.

For the case of anomaly detection, it is desirable to determine where or when an abnormal event was, together with its cause, e.g., what model rules were broken. Hierarchical models are apt for this purpose. They allow to distinguish variables that are more directly related to the observation from the sensors (at the lowest levels of the hierarchy) or more conceptual representations (at the highest levels of the hierarchy). Therefore, this allows distinguishing where the anomaly is located in the hierarchy too. Methods of this type are the Markov Jump Particle Filter (MJPF) [9] and the Rao-blackwellized Particle Filter [10], with the first of the two offering clearer and more reusable semantics than the second one.

Bio-inspired theories as the ones of Friston, Haykin, and Damasio [11–14] have guided the field of self-aware agents. In particular, Friston proposed using Hierarchical Generalized State Filters and introduced linear attractors [11]. Two types of motions are considered: the one in the attractor's direction and the one in the orthogonal direction. The first type of movement can be described as the motivation that the agent pursues and how it moves along the attractor's direction. In contrast, the second type of motion can be connected to the modality with which the agent reaches the attractor along the orthogonal direction, e.g., smoother when the agent is an expert in its task, oscillating when it is uncertain. Therefore, the features related to the two rules of motion can be used to identify a particular behavior and to determine the abnormality of that behavior w.r.t. an unknown one. When multiple behavior models are known, each behavior's discriminatory features can be extracted, and anomalies used for classification. An application example is driver behavior analysis, a field with vast literature [15–17] and one of the main objectives of distinguishing risky and dangerous drivers from safe ones. Oscillating and uncertain drivers fall into the risk category of driving behaviors [15].

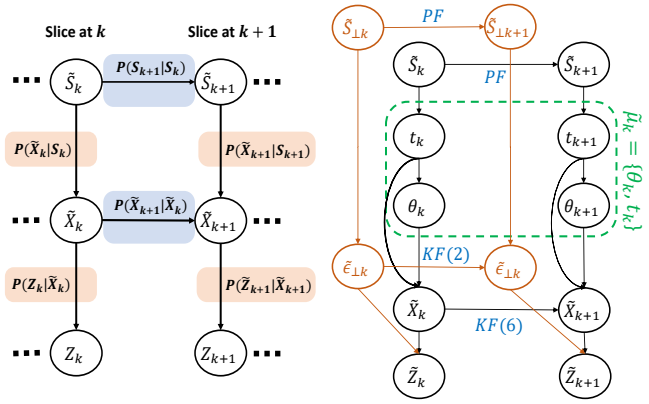
In this paper, we propose an extension of the MJPF presented in [9] to create more precise rules describing the evolution of an agent's state and the objective of treating the study of the evolution along the direction of motion together with the evolution along its orthogonal direction. Tests are conducted on two-dimensional real data. Consequently, the paper's main contributions are the following: *i)* the tracking of parameters at the base of the vehicle's motion and their use to predict the next state. This allows to improve the interpretability of the model, increasing the decomposability; *ii)* the extraction of the features related to the direction perpendicular to motion using the concept of *vorticity* and their usage to define an anomaly related to driver experience/uncertainty; *iii)* the recognition of discriminatory features of a behavior class and the use of the extracted anomalies for behavior classification purposes.

The rest of the paper is organized as follows: Section 2 briefly summarizes related work, Section 3 describes the proposed method, Section 4 discusses the used datasets, and the obtained results, and Section 5 draws the conclusions and suggests future developments.

2. RELATED WORK

The method proposed in this paper is an extension of the one described in [9]. In [9], a DBN architecture was learned from a training dataset. DBNs [7] are a type of Probabilistic Graphical Model (PGM) enabling to discover the causal relationships between variables at consequent time instants (inter-frame dependencies) and at the same time instant (intra-frame dependencies). In particular, MJPF synthesizes a two-level DBN as the one displayed in Fig. 1a : the evolution of the continuous state $\tilde{X}_k$ related to a sensor observation Z_k is tracked on the lower level, whereas discrete variables $\tilde{S}_k$ are used to switch from one linear dynamical model for continuous

978-1-7281-7289-7/20/$31.00 ©2020 IEEE



(a) (b)

Fig. 1: (a): DBN structure of MJPF proposed in [9]. Inter-slice links are shown in orange; temporal-links are colored in blue. (b): DBN structure of proposed G-MJPF.

state prediction to another one. The tracked state $\tilde{X}_k$ is a Generalized State (GS), according to Friston's definition [11]. GSs contain information about the states per se and their higher-order dynamical features, i.e., their motion, velocity, acceleration, etc.. In the MJPF proposed in [9] (from here on referenced as *base MJPF*), the GS is composed by the state and its velocity. The relationship between an observation Z_k and a GS $\tilde{X}_k$ at time instant k is defined as:

$$Z_k = H\tilde{X}_k + \nu_k, \tag{1}$$

where ν_k is assumed to be a zero-mean Gaussian distribution representing the observation noise. The dynamical model, used to perform prediction at the continuous level, is defined as:

$$\tilde{X}_{k+1} = F\tilde{X}_k + BU_{S_k} + \omega_k, \tag{2}$$

where $F\tilde{X}_k$ takes the state-space information from $\tilde{X}_k$ and makes null its time derivatives. BU_{S_k} encodes the time derivative information (actions) of the agent at time k. U_{S_k} depends on the variable S_k, which corresponds to the active cluster at the time k. The variable ω_k is a zero-mean Gaussian distribution representing the noise of the dynamical modeling.

The MJPF uses a set of Kalman Filters (KFs) at the state level, governed by Eq. 1 and by Eq. 2, and a Particle Filter (PF) at the cluster level. It is used to perform anomaly detection on a variety of applications [2, 9, 18]. This paper extends the previous work to build a more flexible and interpretable model based on a four-level DBN. We call this model Generalized-MJPF (G-MJPF).

3. METHOD DESCRIPTION

General architecture. We can divide the description of the method into two parts: *i)* given a first dataset (i.e., a training dataset), we apply on it a Null Force Filter (NFF), which supposes that the tracked object is not affected by any force and continues to move with the same speed. We extract the GSs and the model errors and use them to perform clustering and learn a DBN architecture and a new filter adapted to the dataset. We also extract the prediction error of this filter along the orthogonal direction of motion and build a base MJPF, which allows us to track the information related to how the agent is oscillating around the expected motion. We call the overall obtained model a G-MJPF; *ii)* then, given a second dataset (i.e., a testing dataset), we apply the G-MJPF to detect anomalies w.r.t. the learned model. The description of the method is shown in Fig. 2a. Additionally, we consider using the learned filters and the found anomalies for the application on behavior classification, i.e., we consider the case in which multiple models are present and the recognition of their salient interpretable features.

3.1. Training phase

Generalized States. Let us suppose to be given a training dataset composed of K consequent observations $\{\tilde{Z}_k\}_{k=1\dots K}$ from a sensor. The Generalized Observations (GOs) are composed by the observations (e.g., position data) and their generalized coordinates of motion. For simplicity, we consider as the generalized coordinates of motion the first-time derivative $\dot{Z}_k$ only. Consequently, we can define $\tilde{Z}_k = [Z_k\ \dot{Z}_k]$. Starting from the GOs, we can link GSs and GOs through Eq. 1, where H is an identity matrix. As in [9], we can define the GS as $\tilde{X}_k = [X_k\ \dot{X}_k]^\intercal$.

Null Force Filter. As initial step, we track the evolution of the GSs $\{\tilde{X}_k\}_{k=1\dots K}$, using a NFF, a KF that supposes that the agent is not affected by any force and continues in its motion with unmodified speed w.r.t. the previous time-steps. The dynamic model supposed by the NFF can be expressed through the following equation:

$$\tilde{X}_{k+1} = A\tilde{X}_k + \omega_k, \tag{3}$$

where $A = [A1, A2]$, with $A1 = [I_{d,d}, 0_{d,d}]^\intercal$ and $A2 = [I_{d,d}, I_{d,d}]^\intercal$, being $I_{d,d}$ the identity matrix with d rows and columns, where d is the observation dimension. In the case of the trajectory data, $d = 2$.

Therefore, at each time step in which the NFF is applied, we can extract the desired motion parameters $\tilde{\mu}_k$, which allows us to correct our model, coherently with Friston's free energy principle [12]. Eq. 3 could be corrected as:

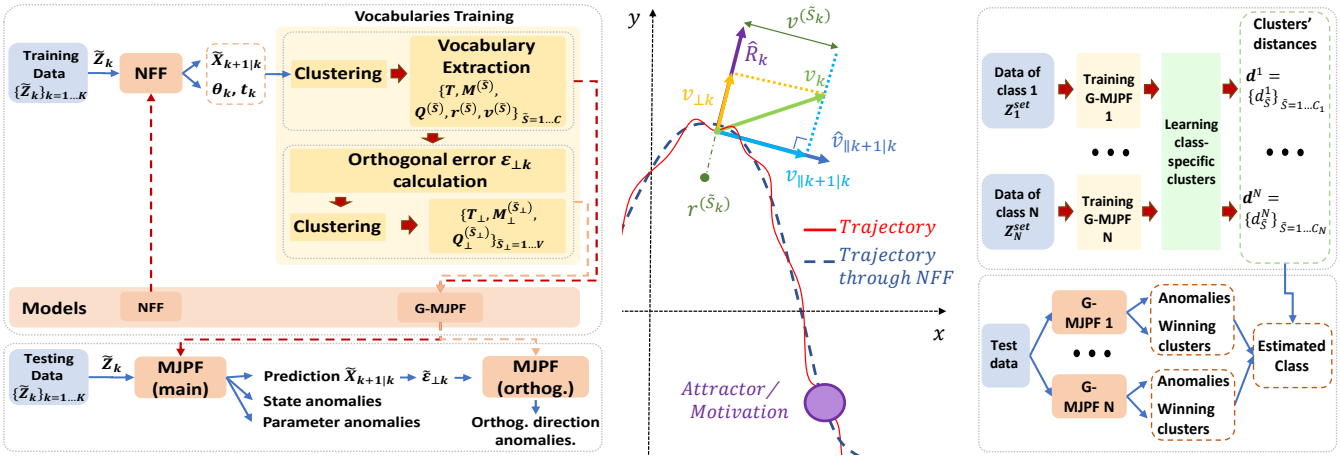
$$\tilde{X}_{k+1} = A\tilde{X}_k + \Phi\tilde{X}_k + \Psi + \omega_k, \tag{4}$$

where Φ represents a rotational correction and Ψ an acceleration correction. Φ can be modeled as $\Phi = [\Phi 1, \Phi 2]$, where $\Phi 1 = 0_{d,d*2}$ and $\Phi 2 = \begin{bmatrix}\phi\\ \phi\end{bmatrix}$, defining ϕ as:

$$\phi = \begin{bmatrix} cos\theta_k - 1 & -sin\theta_k \\ sin\theta_k & cos\theta_k - 1 \end{bmatrix}, \tag{5}$$

being θ_k the rotation angle. Ψ can instead be modeled as $\Psi = [t_k, t_k]^\intercal$, being $t_k = [tx_k, ty_k]$, i.e. the accelerations to add along the d dimensions. We extract the rotation angle first, and then we estimate the acceleration from the remaining error present in the model. In this way, we have extracted the parameters that define our rule of motion over $\tilde{X}_k$, i.e., $\tilde{\mu}_k = \{\theta_k, t_k\}$.

Clustering of GSs and parameters. After performing testing with the NFF, and obtaining the GSs and parameters of motion along the direction of attraction, i.e., $\{\tilde{X}_k, \tilde{\mu}_k\}$, we use the Growing Neural Gas (GNG) [19] algorithm to cluster them. Therefore, clusters group together similar states characterized by similar rules of motion. To each cluster $\tilde{S} = 1 \dots C$ is associated a mean value $M^{(\tilde{S})}$ and a covariance $Q^{(\tilde{S})}$. A transition matrix T describes the probability of transitioning between clusters. Additionally, for each cluster $\tilde{S}$, the rotation center $r^{(\tilde{S})}$ and the mean velocity norm $v^{(\tilde{S})}$ are extracted; $r^{(\tilde{S})}$ being obtained by supposing to move for each $\tilde{X}_k$ using a θ_k equal to the mean θ_k of the cluster. Consequently, the prediction model can be reformulated again: instead of using Eq. 4 to predict how $\tilde{X}_k$ will evolve in the next time instant, the normal to the line between $r^{(\tilde{S})}$ and X_k is found, and $v^{(\tilde{S})}$ is projected along it and summed to X_k. The prediction equation can consequently be reformulated through a non-linear function f written as follows:



(a) (b) (c)

Fig. 2: (a): Training and Testing phases of G-MJPF. (b): Geometric representation of the used variables. (c): Driver behavior classification.

$$\tilde{X}_{k+1} = f(\tilde{X}_k, r^{(\tilde{S})}, v^{(\tilde{S})}) + \omega_k, \tag{6}$$

This model allows considering the possibility of having a primary direction of rotation while not depending on the velocity at the previous time instant, as in Eq. 4, which made the model sensible to noise and less robust for anomaly detection purposes. To note that $r^{(\tilde{S})}$ is extracted after clustering performance and not during it, to avoid a noisy measurement of it. If clusters with similar motions and in close positions happen to have different rotation points, clustering can be refined, considering $r^{(\tilde{S})}$ as additional clustering input.

Extraction of model for orthogonal space. Through the definition of the clusters and their respective type of motion, the features related to the motivation that guides the agent have been extracted, i.e., to find element $v_{\parallel k}$ in Fig. 2b. Based upon [11], also the motion orthogonal to the direction towards the attractor can be modeled. These features are not related to an attractor but rather to the type of agent performing the motion, e.g., how expert or uncertain it is. We define this oscillation in the direction perpendicular to motion with the name of *vorticity*, inspired by the homonym concept used in fluidodynamics.

For each time instant k of training data, based on the assigned clusters, prediction is performed using $r^{(\tilde{S})}$ and $v^{(\tilde{S})}$ as defined in Eq. 6. Consequently, the predicted velocity of the agent towards the attractor is found, i.e., $v_{\parallel k+1|k}$. The error related to this prediction is extracted and projected along the orthogonal direction $\hat{R}_k$ to $v_{\parallel k+1|k}$. We define this orthogonal error as $\epsilon_{\perp k}$. A Generalized Error (GE) is defined from it by considering its first time order derivative $\dot{\epsilon}_{\perp k}$, i.e., $\tilde{\epsilon}_{\perp k} = [\epsilon_{\perp k}, \dot{\epsilon}_{\perp k}]^\intercal$. The geometric representation of the described variables is shown in Fig. 2b.

Clustering using GNG is then performed on the GEs. Consequently, a set of V clusters $\tilde{S}_\perp = 1 \dots V$ is extracted, each associated with a mean value $M_\perp^{\tilde{S}}$ and a covariance $Q_\perp^{(\tilde{S})}$. A transition matrix $T_\perp$ is calculated too. To summarize, we build a model for the orthogonal error using a base MJPF.

3.2. Testing phase

During the testing phase, anomaly detection is performed on a testing dataset. In [9], as described in Section 2, a MJPF was used for the purpose of anomaly detection. In this paper, we propose a modified version of the MJPF, with more precise clustering and general motion rules, allowing the separation and combined tracking of motion along the direction of attraction and along its normal. In the following description, we will consequently concentrate on the differences between the two algorithms.

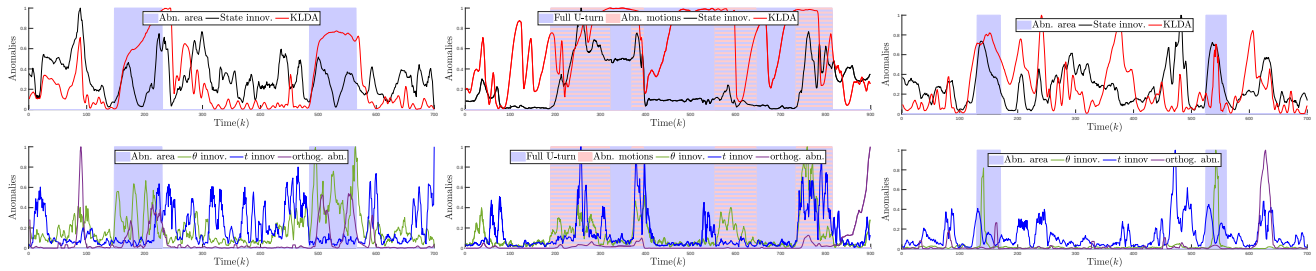
DBN description. During the training phase, the vocabulary for our G-MJPF has been learned; this can be assimilated to the learning of a DBN. Fig. 1b displays the learned DBN: in black, we show the variables related to the motion along the direction to the attractor, highlighting in green the parameters at the base of motion; in orange, we display the variables connected with the normal to the direction towards the attractor; the writings in blue define the filter used to perform prediction at the considered level, referencing the corresponding equation inside the parentheses. To note that the level related to t_k is reported below the one related to θ_k, as we suppose for rotation angle θ_k to be extracted first, and acceleration t_k to be derived as remaining error.

MJPF description. As in the base MJPF, two steps are performed: *prediction* and *update*.

During the prediction phase, at each time instant k, based on the cluster associated to each particle of the Particle Filter (PF), we perform a prediction of the GSs $\tilde{X}_{k+1|k}$ as seen in Eq. 6 for mean value prediction, and the GSs-related rows and columns in $Q^{(\tilde{S})}$ as prediction covariance Q. As in the base MJPF, prediction at the cluster level is performed using the transition matrix T, which is here, however, built through a clustering over both GSs and parameters.

During the update phase, at each time instant $k+1$, the motion parameters θ_{k+1} and t_{k+1} are estimated as in the NFF, and the state prediction $\tilde{X}_{k+1|k}$ is corrected based on the sensor observation, similarly to how performed in [9]. Anomalies are extracted. Additionally, the orthogonal error ϵ_k is found for each particle prediction. The error of the particle with the highest weight is given as input to the parallel MJPF for tracking along the orthogonal direction $\hat{R}_k$. Filtering in this MJPF is performed exactly as in [9].

Anomaly detection. During the update phase of the two parallel MJPFs, anomalies can be extracted on all levels of the hierarchical DBN. Using direct mean subtraction between prediction and update or probabilistic measures based on the Bhattacharya distance as in [9], anomalies on $\tilde{X}_{\parallel k}$, θ_k, t_k and $\tilde{\epsilon}_k$ can be extracted. Using Kullback-Leibler Divergence as in [2], an anomaly at the cluster level can be found. By decomposing the motion along its different directions and parameters, it is now possible to explain the variable at the base of each anomaly signal.



a) b) c)

Fig. 3: Anomalies at state level and cluster level (above) and at parameters level and orthogonal direction level (below) for PA (a), U-turn (b) and ES (c) case.

Table 1: Comparisons between base MJPF [9] and G-MJPF

	Prediction model	Parameters ($\tilde{\mu}_k$) filtering	Orthogonal ($\tilde{\epsilon}_{\perp k}$) filtering	Clustering
Base MJPF [9]	linear	absent	absent	based on GSs only
G-MJPF	rotational	present for improved interpretability	present for improved interpretability	based on GSs & motion parameters

3.3. Behavior classification

The learned model and anomaly detection method can be used for behavior classification, as displayed in Fig. 2c. If we suppose to be provided with N training datasets $Z_1^{set} \dots Z_N^{set}$, one per behavior class, we can learn a G-MJPF on each of them. Found clusters of each set constitute a graph, as observed in [20]. Consequently, we have N Graphs $G_1 \dots G_N$. A Graph Matching procedure can be performed on each graph couple to find which clusters correspond to each other in the two graphs and to detect the clusters that are discriminatory of a class. In this paper, we consider a very simple method to perform this task: when comparing a source graph G_s with one of the $N-1$ targets graphs $G_{t,1} \dots G_{t,i} \dots G_{t,N-1}$, we match the corresponding clusters by finding, for each cluster of G_s, the cluster of $G_{t,i}$ with smallest euclidean distance. Being C_s the number of clusters of G_s, a set of C_s euclidean distances are found. The mean is calculated over the sets obtained from all $G_{t,i}$, finding the normalized cluster distances $\mathbf{d}^s = \{d^s_{\tilde{S}}\}_{\tilde{S}=1\dots C_s}$ of G_s, with $s = 1 \dots N$. The highest distances in the set represent clusters that are more specific to the corresponding source graph, i.e., to the class.

When given a new dataset to perform classification, the N G-MJPFs are applied in parallel on it, and the corresponding anomalies at the different levels are extracted. The sequence of winning clusters is also memorized, i.e., the clusters with the highest weight in the PF at each time k. Each anomaly a_k related to G-MJPF s is modified as $a_k = a_k * (d^s_{\tilde{S}_k}/max(\mathbf{d}^s) * \alpha + \beta)$, where α and β are scalars to weight the impact of the distances on the original anomalies. The use of the cluster's distances allows giving more importance in the classification to those clusters identified as specific of the particular class, providing interpretability also when multiple classes are present. For each G-MJPF, anomalies across levels are normalized and averaged over all time instants. The G-MJPF displaying the lowest final anomaly corresponds to the estimated class.

4. RESULTS

4.1. Dataset description

To test the proposed method, we use different datasets:

ICab data [21]: a dataset from a real vehicle called iCab, performing Perimeter Monitoring (PM) of a closed environment during training, and being hindered by the presence of pedestrians during testing. Testing scenarios include Pedestrian Avoidance (PA), U-turn, and Emergency Stop (ES). Fig. 3 displays the obtained anomalies in the three cases.

UAH-DriveSet dataset [22]: a dataset for driver behavior analysis composed of various car sensory data from six drivers performing two routes (motorway and secondary road) with three types of behaviors (normal, drowsy, and aggressive). In this paper, we used the GPS and accelerometer data of five drivers from the dataset's motorway road. Due to GPS having a sampling rate of 1 Hz only, we used a KF combining GPS and accelerometer data (10 Hz), to obtain a 10 Hz estimation of trajectory data.

4.2. Results description

Anomaly detection on ICab data. We use PM data as training and perform anomaly detection on PA, U-turn, and ES cases. Fig. 3 displays, above, the state and cluster anomalies. To note how state anomalies are noisy, whereas cluster anomalies (KLDA) are better but do not carry specific information about the cause of the anomaly. Using the anomalies on the parameter space, it is now possible to infer that angle of rotation anomalies are present in the avoidance zone in Fig. 3a (in blue), in the U-turn motion, and in the curves in the opposite direction in Fig. 3b (in blue/red). Changes in acceleration generate noisy anomalies at the state level. Additionally, in 3b, the very high anomalies at cluster level in the U-turn zone, not corresponding to rotation or acceleration anomalies, are due to performing motion in the opposite direction to that of training and, therefore, to crossing clusters in an abnormal order. Anomalies on the orthogonal direction are also displayed, corresponding to zones where the vehicle oscillates (e.g., when performing avoidance).

Driver behavior classification on UAH-DriveSet dataset. For each of the considered three driver behavior classes, we perform training of the corresponding G-MJPF and extraction of cluster distances, excluding each time one of the five trajectories. We repeat this for each trajectory. Then, each trajectory is used during testing against the three models that did not include it. To perform classification, we used five anomaly distances on the state along its direction of motion and on the motion parameters, setting $\alpha = 100$ and $\beta = 2$. Obtained accuracy was 73.33%.

5. CONCLUSIONS AND FUTURE WORK

This paper proposes a method to learn prediction models of the state of an object along the direction towards its motivation and along the orthogonal direction. A G-MJPF is developed to perform anomaly detection and is additionally used for driver behavior classification.

Future work will extend the proposed method to higher dimensional data, e.g., combined data from different sensors or video data.

6. REFERENCES

[1] G. Slavic, D. Campo, M. Baydoun, P. Marín, D. Martín, L. Marcenaro, and C. Regazzoni, "Anomaly detection in video data based on probabilistic latent space models," in *IEEE Conference on Evolving and Adaptive Intelligent Systems*, 2020, pp. 1–8.

[2] A. Krayani, M. Baydoun, L. Marcenaro, A. S. Alam, and C. S. Regazzoni, "Self-Learning Bayesian Generative Models for Jammer Detection in Cognitive-UAV-Radios," in *IEEE Global Communications Conference: Cognitive Radio and AI-Enabled Network Symposium*, 2020.

[3] Y. Lu, K. Maheshkumar, S. S. Nabavi, and Y. Wang, "Future frame prediction using convolutional vrnn for anomaly detection," in *IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2019, pp. 1–8.

[4] X. Shi, Z. Chen, H. Wang, D. Yeung, W. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, 2015, pp. 802–810.

[5] L. Zhang, L. Lu, X. Wang, R. Zhu, M. Bagheri, R. M. Summers, and J. Yao, "Spatio-temporal convolutional lstms for tumor growth prediction by learning 4d longitudinal patient data," *IEEE Trans. Medical Imaging*, vol. 39, no. 4, pp. 1114–1126, 2020.

[6] Z. C. Lipton, "The mythos of model interpretability," *Commun. ACM*, vol. 61, no. 10, pp. 36–43, 2018.

[7] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*, The MIT Press, 2009.

[8] Z. Ghahramani, "Learning dynamic bayesian networks," in *Adaptive Processing of Sequences and Data Structures: International Summer School on Neural Networks*. Springer, 1998, pp. 168–197.

[9] M. Baydoun, D. Campo, V. Sanguineti, L. Marcenaro, A. Cavallaro, and C. Regazzoni, "Learning switching models for abnormality detection for autonomous driving," in *International Conference on Information Fusion*, 2018, pp. 2606–2613.

[10] A. Doucet, N. de Freitas, K. P. Murphy, and S. J. Russell, "Rao-blackwellised particle filtering for dynamic bayesian networks," in *Conference in Uncertainty in Artificial Intelligence*, 2000, pp. 176–183.

[11] K. J. Friston, B. S., and G. A., "Cognitive dynamics: From attractors to active inference," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 427–445, 2014.

[12] J. Kilner K. Friston and L. Harrison, "A free energy principle for the brain," *J. Physiol*, vol. 100, no. 1-3, pp. 70–87, 206.

[13] S. Haykin and J. M. Fuster, "On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 608—-628, 2014.

[14] A. R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, Harcourt Brace, 1999.

[15] E. Cheung, A. Bera, E. Kubin, K. Gray, and D. Manocha, "Identifying driver behaviors using trajectory features for vehicle navigation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2018, Madrid, Spain, October 1-5, 2018*. 2018, pp. 3445–3452, IEEE.

[16] J. J. Lu Y. Liu Q. Xue, K. Wang, "Rapid driving style recognition in car-following using machine learning and vehicle trajectory data," *Journal of Advanced Transportation*, vol. 2019, 2019.

[17] M. Brambilla, P. Mascetti, and A. Mauri, "Comparison of different driving style analysis approaches based on trip segmentation over GPS information," in *2017 IEEE International Conference on Big Data*. 2017, pp. 3784–3791, IEEE Computer Society.

[18] D. Kanapram, D. Campo, Mohamad Baydoun, L. Marcenaro, E. Bodanese, C. Regazzoni, and M. Marchese, "Dynamic bayesian approach for decision-making in ego-things," in *IEEE 5th World Forum on Internet of Things*, 2019, pp. 909–914.

[19] B. Fritzke, "A growing neural gas network learns topologies," in *Conference on Neural Information Processing Systems*, 1994, pp. 625–632.

[20] H. Zaal, M. Baydoun, D. Campo, L. Marcenaro, and C. Regazzoni, "Incremental learning of abnormalities in autonomous systems," 2019.

[21] P. Marın-Plaza, J. Beltrán, A. Hussein, B. Musleh, D. Martın, A. de la Escalera, and J. M. Armingol, "Stereo vision-based local occupancy grid map for autonomous navigation in ros," *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pp. 703–708, 2016.

[22] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? presenting the public uah-driveset," in *19th IEEE International Conference on Intelligent Transportation Systems, ITSC*. 2016, pp. 387–392, IEEE.

DELIBERATION FOR INTRA-VEHICLE ROBOTIC ACTIVITIES IN SPACE

Abiola Akanni J. Benton Robert Morris

Intelligent Systems Division
NASA Ames Research Center

ABSTRACT

Intra-Vehicular Robotics (IVR) for space exploration vehicles describes robotic capabilities to perform Intra-vehicle activity (IVA) in an autonomous or remotely operated manner. This paper focuses on autonomy, and more specifically, on the potential application of deliberation functions in robotics to enabling autonomous IVR. We provide an overview of the capabilities required to enable goal-directed operations, robotic systems' ability to autonomously transfer a high-level goal into a set of tasks to accomplish them.

1. INTRODUCTION

Intra-Vehicular Robotics (IVR) for space exploration vehicles refers to robots capable of performing Intra-vehicle activity (IVA) in an autonomous or remotely operated manner. IVAs include state assessment (including inspection, inventory, anomaly detection), logistics (moving and stowing cargo), fault management (all phases), and science operations. NASA researchers explore IVR on Gateway, a spaceport in lunar orbit that will serve as a gateway to deep space and the lunar surface. Since Gateway will primarily be uncrewed for nine to eleven months out of the year, IVR is critical and essential to maintain and protect the vehicle.

This paper focuses on autonomous IVR, and more specifically, on the potential application of deliberation functions in robotics to enabling autonomous IVR. Deliberation functions include planning, acting (refining actions into sensory-motor control), monitoring, observing, and learning. This work has produced an architectural design and closed-loop implementation of a system for goal-directed commanding, automated goal management, task planning, robust execution, execution monitoring, and replanning.

The goal of this paper is to define and illustrate the role of deliberative functions for IVR. The rest of the article is as follows: we define vehicle system management and use the Gateway mission as a working example; then, we define an approach to deliberation architectures, functionality for IVR applications, including extensions for multi-robot deliberation. The work summarized here encapsulates a multi-year effort to demonstrate the effective use of task planning in IVR science activities for robotic manipulators.

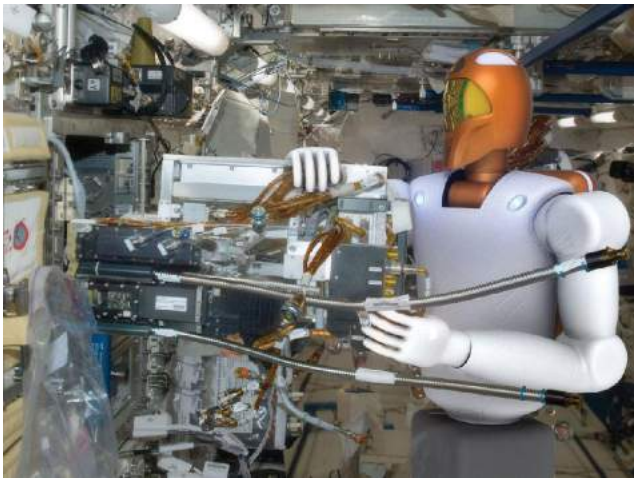


Fig. 1. R2 robotic assistant on the International Space Station

2. EXPLORATION VEHICLE MANAGEMENT

Space vehicle system management is a joint effort of ground systems and personnel, vehicle systems, robotic systems, and vehicle crew. IVR services support all types of operations and vehicle system management. Two vehicle systems management methods are emerging: the first is necessary when a team is on board and operating the vehicle but requires support that the ground controllers cannot give. The second involves vehicle control when no one is on board. Uncrewed vehicle management includes so-called "dormant" periods, when the vehicle is in a state when only the minimal subsystems are required to maintain system health, as well as periods of fully operational autonomous operations [1].

Exploration Missions (EMs) will extend human presence into deep space. Current designs for EMs include incrementally building infrastructure, operational experience, and testing of systems required for long-duration missions in deep space. These activities will be conducted on the Deep Space Gateway (DSG) [2]. DSG is a spaceport in cis-lunar orbit that will serve as a vehicle to deep space and the lunar surface.

Intra-Vehicle Activity (IVA) includes the following:

- *State assessment*, including inspection of vehicle, in-

ventory, and detection of off-nominal conditions;

- *Logistics*, including cargo transport and stowage and opening and closing hatches;
- *Integrated fault management*, including detection, isolation and repair; and
- *Science operations*, for example, biological experiments involving the manipulation and imaging of biological samples.

Each of these activities potentially involves complex, coordinated planning. Furthermore, some activities can be viewed as routine (predictable, performed periodically). In contrast, others might be conducted in response to an off-nominal event, such as detecting a leak or unexplained noise. This suggests the need for *continuous* operations planning, the ability to accept new goals at any time during operations [3].

This work is being applied as part of the Integrated System for Autonomous and Adaptive Care-taking (ISAAC) program, a software system to monitor the telemetry from the International Space Station (ISS) systems, and, eventually, on the DSG. The goal of the ISAAC program is to provide autonomous spacecraft caretaking during uncrewed periods.

To illustrate the technical challenges of IVR for EMs, use cases involving cargo transport logistics and integrated fault management are evaluated. The following sections summarize these use cases.

2.0.1. Example: Transporting Cargo Bags

In this scenario, one or more Cargo Transfer Bags (CTBs) need to be retrieved and transported between a cargo vehicle and a Gateway module. A typical logistics task might require the robot to approach and grasp a CTB with a magnetic gripper. The CTB is equipped with a bag fixture that enables magnetic gripping, and the stowage location is equipped with a berth fixture to hold the item.

Following the grasp, the bag is magnetically released from its stowage location. The robot transports the CTB to its targeted location on Gateway, at which time it is stowed on a new stowage location, using a similar sequence involves magnetic connectors. In addition to the routine grasp/ungrasp/transport actions, solving logistics can also involve set-up tasks such as installing bag fixtures or berth fixtures. Logistics may involve the coordination of multiple robots performing different subtasks.

2.0.2. Example: Integrated Fault Management

In this scenario, a micro-meteoroid strike has caused a leak in a Gateway module during an uncrewed mission phase. The leak must be patched within a matter of hours to avoid significant impacts, such as losing pressurized payloads that cannot tolerate depressurization.

There are three leak management phases: detection, localization, and mitigation. During the detection phase, vehicle sensor information such as pressure sensor trend analysis or the sound of thrust generated by the leak is used to signal the presence of a leak.

The second phase is localization, which works through several mechanisms. It uses coarse localization, where coarse sensor data is used to localize within, say, a 2 by 2-meter area. It begins with a survey procedure, requiring preparation actions like turning off noisy systems that mask the leak noise. We then perform preparation, where a mobile inspection robot with ultrasound sensors can find noise sources indicating the leak. Finally, during report and confirmation, the robot communicates precise leak location and may confirm using other sensor data.

During the mitigation phase, and depending on the location and type of leak, a mobile manipulator robot may patch it using a patch kit. Otherwise, mitigation would focus on steps like moving sensitive equipment out of the affected module and closing hatches to isolate it from the rest of the vehicle.

The examples of logistics and integrated fault management illustrate several technical challenges for effective autonomous operations. Among those challenges are:

- at least some of the actions may be stochastic: for example, the result or effect of a sensing action may not be known with certainty; furthermore, the duration, as well as the success of an activity (e.g., surveying an area), might not be determined;
- the activity may be time-critical; a sequence of actions must be performed before a deadline.
- the state of the world is only partially known at the time that decisions need to be made about what to do; and
- coordination of knowledge and actions is required to attack and solve the problem. This coordination may be robot-robot, or robot-ground system, or robot-vehicle.

Solving these challenges autonomously will require deliberative decision-making capabilities on the part of robots and other autonomous systems. In the next section, we provide an overview of these capabilities.

3. DELIBERATION FOR AUTONOMOUS IVR

Deliberative decision-making can be regarded as a cognitive process resulting in selecting a course of action among several alternative scenarios [4]. Deliberation is a cognitive capability of humans that a machine can automate.

Not all actions require deliberation; some actions can be viewed as automatic or 'purely reactive.' For example, the robotic arm action of grasping a test tube from a holder does not involve deliberation. On the other hand, if the grasp action fails for some reason, then some deliberation might be useful

to decide what to do next (as opposed to merely halting the action). In this case, deliberation can be said to be tightly integrated with the action. In other cases, there is a more apparent separation between deliberation and action. For example, if a robot is assigned a goal to find a leak, there needs to be deliberation (planning) to develop and execute a plan to accomplish the goal. Depending on how deliberation is distributed among humans and machines, the planner may be human, an automated system residing on a separate machine, or part of the deliberation capabilities on the robot itself.

An *autonomous* robot is one in which at least some of the machine's actions are performed without direct external control (with no teleoperation). An autonomous robotic deliberation system [4] is an autonomous system that allows the robot to decide its course of action. Deliberation functions include planning (turning goals into actions), acting (changing the state of the world or an internal state), monitoring (either its actions or events in the world), observing, and learning. Autonomous robotic deliberation has advantages for rich environments in which teleoperation is difficult or impossible. As illustrated in the previous section, this to be the case for a lot of IVR activity.

Deliberation is never a strictly isolated individual activity. For IVA, it is necessary to consider a number of robot interfaces for deliberation:

- **ground-robot**: tele-operation vs goal-based operations [5]
- **crew-robot**: human-robot coordination; for example, robot assistants in space [6]
- **vehicle-robot**: as part of vehicle system/subsystem autonomy [1]
- **robot-robot**: multi-robot coordination [7]

It would be too ambitious here to summarize and survey the past and current work in developing effective interfaces between deliberative robots and external systems. Instead, in the remainder of this paper we focus on deliberative technologies for multi-robot operations (robot-robot deliberative systems).

4. ARCHITECTURES FOR GOAL-BASED OPERATIONS

Most architectures for deliberation systems to enable goal-directed behaviors in robots consist of layers (commonly three) [8]. For example, in previous work by the same authors [9], a layered architecture implemented goal-directed behavior for a manipulator science assistant. Here we briefly review the principles of a layered approach and the central component systems. Figure 2 illustrates the layered architecture concept with an added component for deliberative coordination between robotic systems.

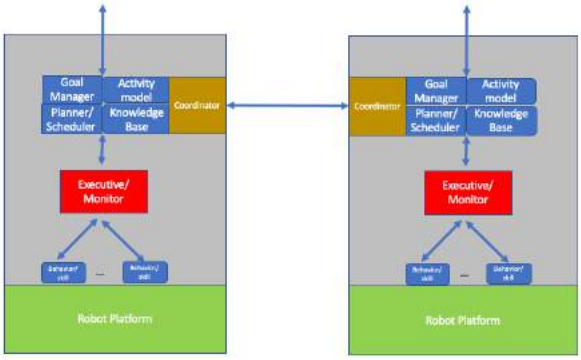


Fig. 2. Layered Architectures for Multiple Robot Deliberation Systems

A goal is a high-level description of the desired result of some activity. IVA is a continuous activity, so new goals emerge at any time, including while other activities are being performed. The system needs to maintain a prioritized list of current goals, which is updated when new goals arrive [10]. This *goal manager* is used both as a pre-processing stage before planning and during execution.

A goal is synthesized into a collection of actions that accomplish it. A plan model defines actions in terms of changes to the world state made by the actions. The Plan Definition Description Language (PDDL) is often used to create these models [11]. Alternatively, timeline-based planning approaches [12] view actions as temporal intervals and planning works by synthesizing a partial plan until all constraints defined on the plan are satisfied.

Executing a plan [13] consists of a set of refinements of actions into commands to the robot platform to transform some part of the world, including the platform itself [14]. Many different approaches to integrating task planning and other high-level deliberation with motion planning, observation systems, and closed-loop actuation have been proposed (for example, [15], [16]).

Execution monitoring is a capability for tracking the world state's evolution while a plan is executed [17], [18]. Systems robust to plan failure should respond to unpredicted changes to the world state that cause a plan to fail by triggering corrective actions.

Finally, added deliberation features are required for several robotic agents to plan and execute together in a shared environment to accomplish a set of common goals [19]. Agents must both synthesize a plan and coordinate with other agents to build a joint plan. Joint planning potentially involves communicating local plans and knowledge between planning or execution components.

5. IMPLEMENTATION

Our approach to implementing deliberation capabilities expands upon the work previously conducted by the same team [9] by extending them to a multiple robot scenario using a decentralized approach, as depicted in Figure 2. Currently the focus is on deliberation systems for a free flyer (Astrobee [20]) and a dexterous manipulator (R2, as shown in Figure 1) to solve logistics and fault management tasks described above. The architecture for deliberation is being implemented using ROSPlan [21], which is being extended for multi-robot operations. ROSPlan integrates both software and hardware solution to streamline the process of task execution, robot cooperation, and task re-planning. A PDDL multi-agent planner that plugs into ROSPlan will interface with kinetics and motion planning.

To summarize, goal-based autonomous operations will automate the process of transforming high-level tasks into a coordinated plan that the robots will jointly execute. An execution monitoring system will detect, classify and recover from plan failures, thus ensuring robustness from the uncertainties confronted by operating in a dynamic environment.

6. SUMMARY

This paper has provided an overview of deliberation systems for Intra-Vehicle Robotic systems on future exploration vehicles. These ideas are currently being implemented to enable robots to solve logistics problems in simulation and be soon used in missions.

7. ACKNOWLEDGEMENTS

This work was performed under NASA's Integrated System for Autonomous and Adaptive Care-taking (ISAAC) Project.

8. REFERENCES

[1] J. M. Badger, P. Strawser, and C. Claunch, "A distributed hierarchical framework for autonomous spacecraft control," in *2019 IEEE Aerospace Conference*, 2019, pp. 1–8.

[2] Gateway, *Q&A: NASA's New Spaceship*, 2018, https://www.nasa.gov/feature/questions-nasas-new-spaceship.

[3] M. Brenner and Bernhard Nebel, "Continual planning and acting in dynamic multiagent environments," *Journal of Autonomous Agents and Multiagent Systems*, vol. 19, pp. 297–331, 2009.

[4] Felix Ingrand and Malik Ghallab, "Deliberation for autonomous robots," *Artif. Intell.*, vol. 247, no. C, pp. 10–44, June 2017.

[5] Daniel D. Dvorak, Michel D. Ingham, J. Richard Morris, and John Gersh, "Goal-based operations: An overview," *Journal of Aerospace Computing, Information and Communication*, 2012.

[6] Lulu Chang and Trevor Mogg, "Spacex delivers CIMON, along with berries and ice cream, at ISS," in *Emerging Tech. Digital Trends*, 2018.

[7] Zhi Yan, Nicolas Jouandeau, and Arab Ali Cherif, "A survey and analysis of multi-robot coordination," *International Journal of Advanced Robotic Systems*, vol. 10, no. 12, pp. 399, 2013.

[8] Erann Gat, R. Peter Bonnasso, Robin Murphy, and Aaai Press, "On three-layer architectures," in *Artificial Intelligence and Mobile Robots*. 1997, pp. 195–210, AAAI Press.

[9] Shaun Azimi, Emma Zemler, and Robert A. Morris, "Autonomous robotics manipulation for in-space intravehicle activity," in *Workshop on Planning and Robotics (PlanRob)*, 2019.

[10] Swaroop Vattam, Matthew Klenk, Matthew Molineaux, and David W Aha, "Breadth of approaches to goal reasoning: A research survey," Tech. Rep., Naval Research Lab Washington DC, 2013.

[11] M. Cashmore, M. Fox, T. Larkworthy, D. Long, and D. Magazzeni, "Planning inspection tasks for AUVs," in *2013 OCEANS - San Diego*, 2013, pp. 1–8.

[12] Jeremy Frank and Ari Jónsson, "Constraint-based attribute and interval planning," *Constraints*, vol. 8, no. 4, pp. 339–364, 2003.

[13] Phil Kim, Brian C Williams, and Mark Abramson, "Executing reactive, model-based programs through graph-based temporal planning," in *IJCAI*, 2001, pp. 487–493.

[14] Christian Muise, J. Christopher Beck, and Sheila A. McIlraith, "Flexible execution of partial order plans with temporal constraints," *International Joint Conference on Artificial Intelligence*, pp. 2328–2335, 2013.

[15] J. A. Bagnell, F. Cavalcanti, L. Cui, T. Galluzzo, M. Hebert, M. Kazemi, M. Klingensmith, J. Libby, T. Y. Liu, N. Pollard, M. Pivtoraiko, J. Valois, and R. Zhu, "An integrated system for autonomous robotics manipulation," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 2955–2962.

[16] S. Srivastava, E. Fang, L. Riano, R. Chitnis, S. Russell, and P. Abbeel, "Combined task and motion planning through an extensible planner-independent interface layer," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014.

[17] Ola Pettersson, "Execution monitoring in robotics: A survey," *Robotics and Autonomous Systems*, vol. 53, no. 2, pp. 73–88, 2005.

[18] Maria Fox, Alfonso Gerevini, Derek Long, and Ivan Serina, "Plan stability: Replanning versus plan repair," in *ICAPS*, 2006, vol. 6, pp. 212–221.

[19] Alejandro Torreño, Eva Onaindia, Antonín Komenda, and Michal Stolba, "Cooperative multi-agent planning: A survey," *CoRR*, vol. abs/1711.09057, 2017.

[20] Maria G. Bualat, Trey Smith, Ernest E. Smith, Terrence Fong, DW Wheeler, and the Astrobee Team, "Astrobee: A new tool for ISS operations," in *Proc. SpaceOps (AIAA 2018-2517)*, 2018.

[21] Michael Cashmore, Maria Fox, Derek Long, Daniele Magazzeni, Bram Ridder, Arnau Carrera, Narcis Palomeras, Natalia Huros, and Marc Carreras, "ROSPlan: Planning in the robot operating system," in *Proceedings of the Twenty-Fifth International Conference on Automated Planning and Scheduling*. 2015, pp. 333–341, ICAPS.

EXPERIMENTAL VALIDATION OF DOMAIN KNOWLEDGE ASSISTED ROBOTIC EXPLORATION AND SOURCE LOCALIZATION

Thomas Wiedemann, Dmitriy Shutin

Institute of Communications and Navigation
German Aerospace Center
Oberpfaffenhofen, Germany

Achim J. Lilienthal

Center of Applied Autonomous Sensor Systems
Örebro University
Örebro, Sweden

ABSTRACT

In situations where toxic or dangerous airborne material is leaking, mobile robots equipped with gas sensors are a safe alternative to human reconnaissance. This work presents the Domain Knowledge Assisted Robotic Exploration and Source Localization (DARES) approach. It allows a multi-robot system to localize multiple sources or leaks autonomously and independently of a human operator. The probabilistic approach builds upon domain knowledge in the form of a physical model of gas dispersion and the *a priori* assumption that the dispersion process is driven by multiple but sparsely distributed sources. A formal criterion is used to guide the robots to informative measurement locations and enables inference of the source distribution based on gas concentration measurements. Small-scale indoor experiments under controlled conditions are presented to validate the approach. In all three experiments, three rovers successfully localized two ethanol sources.

Index Terms— mobile robot olfaction, gas source localization, Bayesian inference, swarm exploration.

1. INTRODUCTION

Mobile robotic platforms, like rovers and Unmanned Aerial Vehicles (UAVs), are the means of choice when it comes to exploration missions in hazardous environments. For example, in disaster relief scenarios or Chemical, Biological, Radiological and Nuclear (CBRN) events, mobile robots can be dispatched to survey an area of interest and provide an overview of the current situation. There it is important to explore the environment as fast as possible and provide reliable information early enough to civil protection agencies and first responders. It is obvious that while a single robot would need a certain time to explore a region of interest, multiple robots can accomplish the same task faster.

However, the deployment of multiple robots brings along several challenges. Whereas a single robot can be easily teleoperated, steering and coordinating many robots in real-time is a too complex task for a single operator or even for a team of coordinators. In addition, in disaster scenarios, very few human resources are available for controlling robots. Autonomy, on the other hand, allows to address these challenges: through cooperation, the robots can coordinate themselves and accomplish the exploration task independently of an operator. One of the key elements required to implement such an autonomous multi-robot system is an exploration strategy – an algorithm that allows the robots to decide where to collect information or measurements.

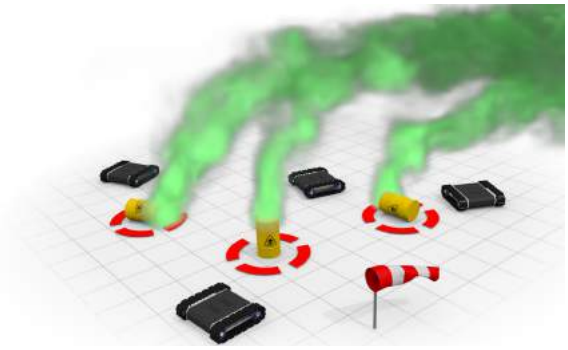


Fig. 1: Robotic gas source localization scenario

In this paper, we consider the task of exploring the dispersion of a toxic or dangerous airborne trace substance (referred to as "gas" in the latter) leaking from an unknown number of sources (see Fig. 1). Our goal is to localize the sources using gas concentration measurements taken by in-situ sensors mounted on the robots. This paper shortly presents our Domain Knowledge Assisted Robotic Exploration and Source Localization (DARES) strategy developed in [1]. As the main contribution, the paper presents results of an evaluation of the DARES approach in experiments under laboratory conditions. In contrast, our previous work studied the approach only in simulations.

2. DOMAIN KNOWLEDGE ASSISTED ROBOTIC EXPLORATION AND SOURCE LOCALIZATION

In the past, many gas source localization strategies for robotic applications were based on the idea that the gas concentra-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

tion rises monotonously with proximity to a source. These approaches are often referred to as *chemotaxis* [2, 3]. Using *chemotaxis* a robot tries to follow the gradient of the gas concentration. Such approaches are often supported by additional information like airflow or wind [4]. However, the monotonicity of the concentration distribution does not hold in many real-world environments, since gas dispersion is disturbed by turbulence [5]. Over time, more sophisticated gas source localization strategies have emerged. They take into account more complex mathematical models of the gas dispersion process [6, 7, 8]. These strategies aim at maximizing the information gain obtained with collected measurement data. Consequently, these methods are termed *infotaxis* [9]. The proposed DARES approach likewise follows an infotactic concept.

Our key idea is to assist the robots by *a priori* available domain knowledge about the gas dispersion process. With this additional information, the robots can localize the sources faster, i.e. with fewer measurements, as shown in [10]. In particular, we assist the robots by providing a physical description of gas dispersion in terms of a Partial Differential Equation (PDE). Additionally, since the exact number of sources is assumed as unknown, we endow the model with an assumption that the sources are sparsely distributed. This weak assumption turned out to be very beneficial in order to localize the sources [1].

In what follows we explain how to encode our knowledge and assumptions in a probabilistic gas dispersion model suitable for estimating the sources from concentration measurements. Afterward, the model is used to design an exploration strategy that guides the robots to informative measurement locations with the objective to reduce the uncertainty of the estimates.

2.1. Probabilistic Gas Dispersion Model

From physics it is known that the gas dispersion process over some domain of interest Ω can be approximated by the (stationary) advection-diffusion PDE [11]:

$$-\nabla^2 f(\boldsymbol{x}) + \boldsymbol{v}(\boldsymbol{x})\,\nabla f(\boldsymbol{x}) = u(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega \tag{1}$$

$$\text{s.t.} \quad f(\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in \partial\Omega \tag{2}$$

where $f(\boldsymbol{x}) = 0$ is a boundary condition on the boundary $\partial\Omega$.

Here, we restrict ourselves to the static two dimensional case ($\Omega \subset \mathbb{R}^2$), where function $f(\boldsymbol{x})$ denotes the gas concentration at location $\boldsymbol{x}$. The right-hand side of (1) models the source distribution. More precisely, the function $u(\boldsymbol{x})$ represents the source strength or amount of material inflow at a location $\boldsymbol{x}$. Furthermore, the vector-valued functions $\boldsymbol{v}(\boldsymbol{x}) \in \mathbb{R}^2$ describes the two components of the airflow field at a location $\boldsymbol{x}$.

However, estimation of a function $u(\boldsymbol{x})$, which in general requires application of calculus of variations, cannot be solved analytically. Instead we approximate (1) numerically using Finite Element Method (FEM) [1]. To this end, the continuous environment Ω is discretized using a finite number of N nodes spanning a mesh. The continuous functions $f, u, \boldsymbol{v}$ are approximated by a finite number of linear shape functions – finite elements – that linearly interpolate between the mesh nodes.[1]. Thus the numerical approximations of the continuous functions can be fully parameterized by the values at the mesh nodes. The values at the mesh nodes are aggregated in vectors $\boldsymbol{f}, \boldsymbol{u}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2 \in \mathbb{R}^N$ corresponding to discretization of the continuous functions f, u, and $\boldsymbol{v}$, respectively. Thus, the variational problem (1) can be equivalently represented with a system of $N + B$ algebraic equations:

$$\begin{cases} r_i(\boldsymbol{f}, \boldsymbol{u}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2) = 0, & i = 1, \ldots, N \\ r_i(\boldsymbol{f}) = 0, & i = N+1, \ldots, N+B \end{cases} \tag{3}$$

where $r_i(\boldsymbol{f}) = 0$, $i = N+1, \ldots, N+B$, represent B equations obtained after discretization of the boundary condition (2) at $\partial\Omega$. Essentially, B is the number for of nodes in the mesh that are located at the border of Ω.

We now cast the deterministic model (3) into a probabilistic setting. Instead of demanding r_i to be exactly zero, we assume that the equations only hold with a certain precision τ_s. The derivations (or residuals) of individual equations are assumed to be spatially and temporally white zero-mean Gaussian samples. This assumption results in the following conditional probability density function for the gas concentration given the source distribution and airflow field:

$$p(\boldsymbol{f}|\boldsymbol{u}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2) \propto \prod_{i=1}^{N+B} \mathrm{e}^{-\frac{\tau_s}{2}(\boldsymbol{f}, \boldsymbol{u}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2)^2}. \tag{4}$$

Furthermore, we assume the L robots to be equipped with in-situ gas sensors. Each robot takes K_l noisy concentration measurements at different locations. Thus, the likelihood function of our model can be defined as:

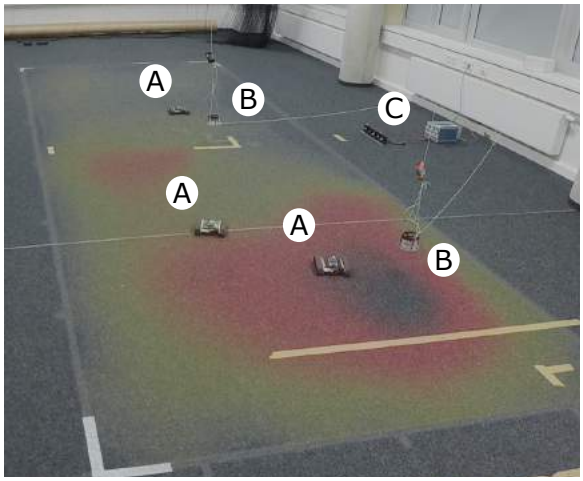
$$p(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L|\boldsymbol{f}) \propto \prod_{l=1}^{L} \exp\left(-\frac{\tau_m}{2}\|\boldsymbol{M}_l^T \boldsymbol{f} - \boldsymbol{y}_l\|^2\right), \tag{5}$$

where $\boldsymbol{y}_l \in \mathbb{R}^{K_l}$, $l = 1, \ldots, L$, are measurements taken by the robot l, $\boldsymbol{M}_l \in \{0, 1\}^{K_l \times N}$ is a binary selection matrix that "picks" elements in $\boldsymbol{f}$ corresponding to the location of the robot when the measurement was taken. Besides, τ_m is the sensor measurement noise precision.

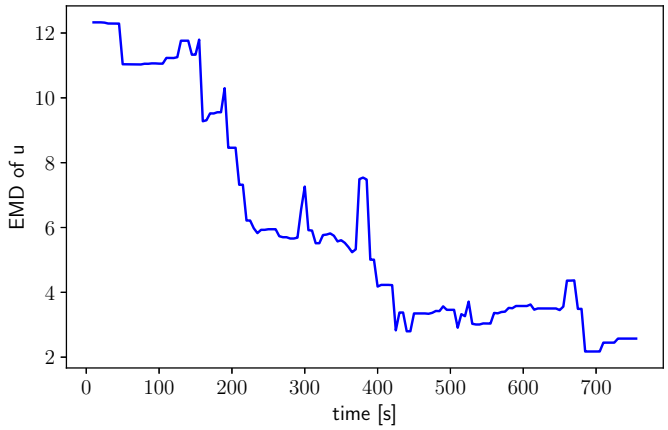
We aim at a Bayesian inference approach, and therefore we also need to define the prior distribution for the airflow $p(\boldsymbol{v}_1, \boldsymbol{v}_2)$, as well as the source prior distribution $p(\boldsymbol{u})$. What concerns the airflow prior, in this work we assume it to be a multivariate Gaussian:

$$p(\boldsymbol{v}_1, \boldsymbol{v}_2) = p(\boldsymbol{v}_1)p(\boldsymbol{v}_2) = N(\boldsymbol{v}_1|\boldsymbol{\mu}_{v1}, \Sigma_{v1})N(\boldsymbol{v}_2|\boldsymbol{\mu}_{v2}, \Sigma_{v2})$$

[1] These are also known as Lagrange elements of first-order



(a)



(b)

Fig. 2: The figure shows in (a) the experimental setup in the laboratory. The robots are indicated with A. The two sources are hanging from the ceiling (indicated with B) so that the robots can drive below without collision. The artificial airflow is generated by fan C. As an overlay, the estimated gas concentration after $12min$ in one experimental run is shown. The error in the estimated source distribution is plotted in (b) over time and averaged over three experiments.

where $\boldsymbol{\mu}_{vj}$, Σ_{vj}, $j = 1, 2$ are fixed design parameters. Practically, these can be set from, e.g., weather forecast or determined using anemometer sensors in the field. The modeling of a source prior is a bit more involved. To incorporate the sparsity assumption we use Sparse Bayesian Learning and represent $p(\boldsymbol{u})$ using a hierarchical prior. The model is augmented with hyperparameters $\boldsymbol{\alpha}$, such that $p(\boldsymbol{u}, \boldsymbol{\alpha}) = p(\boldsymbol{u}|\boldsymbol{\alpha})p(\boldsymbol{\alpha}) = N(\boldsymbol{u}|\mathbf{0}, \mathrm{diag}\{\boldsymbol{\alpha}\}^{-1})\mathrm{Ga}(\boldsymbol{\alpha})$, where $\mathrm{Ga}(\boldsymbol{\alpha})$ is a Probability Density Function (PDF) of a Gamma distribution. The hyperparameters α are estimated alongside other parameters. Based on equation (4) and (5) our Bayesian inference approach gives us the posterior

$$p(\boldsymbol{f}, \boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{\alpha}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L) \propto p(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L|\boldsymbol{f})p(\boldsymbol{f}|\boldsymbol{u}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2)p(\hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2)p(\boldsymbol{u}|\boldsymbol{\alpha})p(\boldsymbol{\alpha}). \quad (6)$$

To maximize this posterior we represent it using a factor graph and perform inference using message passing, which can also be implemented in a distributed setting over the network of robots. Due to space constraints, we refer the reader to our works [12, 10] where the inference algorithm is described in more detail. In our setup, each robot calculates messages of a partition of the whole graph and shares the results with the other robots.

2.2. Exploration Strategy

The maximum of the posterior (6) provide us the source distribution $\boldsymbol{u}$ based on gas concentration measurements $\boldsymbol{y}_l$, $l = 1, \ldots, L$, taken by robots. In order to achieve a high level of autonomy, however, an intelligent sampling strategy is needed that will guide robots to new, informative sampling locations. For this purpose, we propose an uncertainty-driven exploration strategy, where the new measurements are taken at locations that maximally reduce the uncertainty of the obtained estimates. To this end, a gauge for the spatial uncertainty is needed with respect to different locations in the environment.

Here the probabilistic inference approach becomes useful. Recall, that $\boldsymbol{f}$ represents a concentration value at each discretized location of the environment Ω. We can calculate a marginal distribution for each entry f_i, $i = 1, \ldots, N$, of $\boldsymbol{f}$:

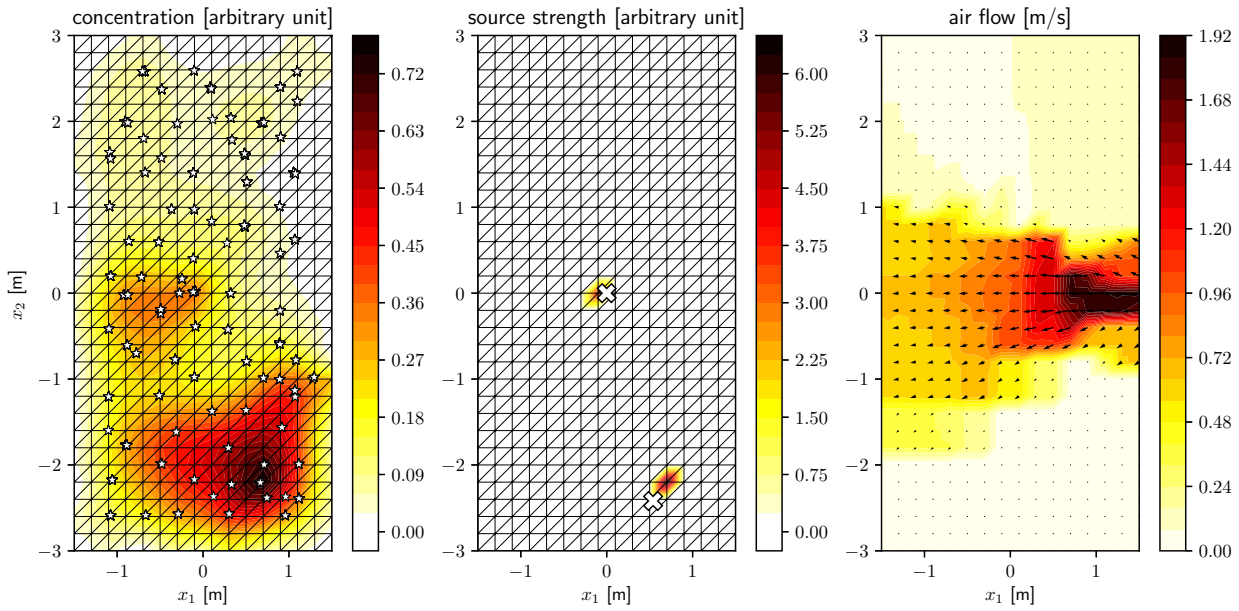
$$p(f_i) \propto \int \cdots \int p(\boldsymbol{f}, \boldsymbol{u}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, \boldsymbol{\alpha}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L) d \sim f_i \approx N(f_i|\mu_{f_i}, \sigma^2_{f_i}); i = 1, \ldots, N \quad (7)$$

where we use the notation $\int \cdots \int d \sim f_i$ to indicate a marginalization operation over all variables except for f_i. The proposed idea approximates $p(f_i)$ with a Gaussian distribution and uses the variance $\sigma^2_{f_i}$ as a gauge of concentration uncertainty.

Robots are then sent to locations with the highest variance, i.e. with the highest uncertainty, which in turn implies the highest Shannon entropy of the concentration value. Taking a measurement at this location would reduce the total concentration uncertainty the most.

3. EXPERIMENTAL EVALUATION

We evaluated our approach in experiments under laboratory conditions as shown in Figure 2. We placed two Petri dishes hanging from the ceiling and filled with ethanol modeling 2 sources. They acted as sources of ethanol vapor in our



(a) (b) (c)

Fig. 3: The figure depicts a snapshot of an exploration experiment after $12min$ (corresponding to Figure 2). In (a) the estimated gas concentration is shown. Besides, the white stars indicate all measurement locations. In (b) the estimated source distribution (spatial source strength) is plotted, where the white crosses indicate the actual position of the ethanol sources. In (c) the artificial airflow field generated by fans and used in the inference approach is shown.

experiments. Above the culture dishes, fans were mounted blowing air downwards to accelerate evaporation and dispersion. Three small robots were deployed in the experiment. The robots were equipped with Photoionization detectors (alphasense PID-AH2) to measure the ethanol concentration in the room. Also, the robots make use of a camera tracking system for precise indoor localization and can move to desired waypoints enabled by a robotic path planner. We generated an artificial airflow in the room using multiple fans. The airflow turned out to be long-term stable and was sampled before the experiments using multiple anemometers. The measured spatial airflow field, as depicted in Figure 3c, was used as the airflow prior in the inference approach (6). Namely, based on the measurements the parameters $\boldsymbol{\mu}_{vj}$, Σ_{vj}, $j = 1, 2$, of the airflow prior were computed. To evaluate the exploration strategy, we compare the estimated source distribution $\boldsymbol{u}$ to the ground truth source distribution $\boldsymbol{u}_{gt}$ by means of the Earth Mover's Distance (EMD) [13]. Vector $\boldsymbol{u}_{gt}$ is an all-zero vector, with an exception of locations corresponding to sources. Note that in reality, the actual source inflow rate is unknown. Therefore, we set the elements in $\boldsymbol{u}_{gt}$ corresponding to sources to 1 and normalized the estimated source distribution to 1, too, before comparison. The resulting EMD performance is plotted in Figure 2b and averaged over three experimental runs. It can be seen that by intelligently sampling the gas concentration the robots successively reduce the error in the estimated source distribution. Figure 3b also shows the estimated source distribution after $12min$ for one experiment. There, the two peaks in the estimated source distribution nearly perfectly match the actual position of the ethanol sources indicted by the two crosses.

4. CONCLUSION

The paper shortly summarizes the DARES approach towards a exploration of gas sources. Based on a physical model of gas dispersion and the assumption that the dispersion process is driven by multiple sparsely distributed sources, a probabilistic model was formulated. The model is the foundation of a Bayesian inference approach to estimate the source distribution based on gas concentration measurements taken by robots. Further, the model is used to derive an exploration strategy that autonomously guides the robots to informative measurement locations. The DARES approach has been evaluated in small-scale experiments in an indoor environment under controlled conditions. There it has been shown that the robots are able to successfully localize ethanol vapor sources. In the future, the DARES approach needs to be evaluated also in a more challenging environment, for example in outdoor scenarios.

5. REFERENCES

[1] Thomas Wiedemann, Achim J. Lilienthal, and Dmitriy Shutin, "Analysis of model mismatch effects for a model-based gas source localization strategy incorporating advection knowledge," *Sensors*, vol. 19, no. 3, 2019.

[2] R. Andrew Russell, Alireza Bab-Hadiashar, Rod L. Shepherd, and Gordon G. Wallace, "A comparison of reactive robot chemotaxis algorithms," *Robotics and Autonomous Systems*, vol. 45, pp. 83–97, 2003.

[3] Gideon Kowadlo and R. Andrew Russell, "Robot odor localization: A taxonomy and survey," *International Journal of Robotics Research*, vol. 27, no. 8, pp. 869–894, 2008.

[4] David J. Harvey, Tien Fu Lu, and Michael A. Keller, "Comparing insect-inspired chemical plume tracking algorithms using a mobile robot," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 307–317, 2008.

[5] Ali Marjovi and Lino Marques, "Multi-robot odor distribution mapping in realistic time-variant conditions," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 3720–3727.

[6] Julian Ruddick, Ali Marjovi, Faezeh Rahbar, and Alcherio Martinoli, "Design and Performance Evaluation of an Infotaxis-Based Three-Dimensional Algorithm for Odor Source Localization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 1413–1420.

[7] Qing Hao Meng, Wei Xing Yang, Yang Wang, and Ming Zeng, "Collective odor source estimation and search in time-variant airflow environments using mobile robots," *Sensors*, vol. 11, pp. 10415–10443, 2011.

[8] Hadi Hajieghrary, M. Ani Hsieh, and Ira B. Schwartz, "Multi-agent search for source localization in a turbulent medium," *Physics Letters, Section A*, vol. 380, no. 20, pp. 1698–1705, 2016.

[9] Massimo Vergassola, Emmanuel Villermaux, and Boris I. Shraiman, "'Infotaxis' as a strategy for searching without gradients," *Nature*, vol. 445, no. 7126, pp. 406–409, 2007.

[10] Thomas Wiedemann, Dmitriy Shutin, and Achim J. Lilienthal, "Model-based gas source localization strategy for a cooperative multi-robot system - A probabilistic approach and experimental validation incorporating physical knowledge and model uncertainties," *Journal of Robotics and Autonomous Systems*, vol. 118, pp. 66–79, 2019.

[11] John Crank, *The Mathematics Of Diffusion*, Clarendon Press, Oxford, second edition, 1975.

[12] Thomas Wiedemann, Christoph Manss, and Dmitriy Shutin, "Multi-agent exploration of spatial dynamical processes under sparsity constraints," *Autonomous Agents and Multi-Agent Systems*, vol. 32, no. 1, pp. 134–162, 2018.

[13] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas, "A Metric for Distributions with Applications to Image Databases," in *IEEE International Conference on Computer Vision*, 1998, pp. 59–66.

A VISION-BASED METHOD FOR ESTIMATING CONTACT FORCES IN INTRACARDIAC CATHETERS

Hamidreza Khodashenas[1]*, Pedram Fekri*[1]*, Mehrdad Zadeh*[2] *and Javad Dargahi*[1]

[1]Mechanical Engineering Department, Concordia University, Montreal, QC, Canada
[2]Electrical and Computer Engineering Department, Kettering University, Flint, Michigan, USA

ABSTRACT

Atrial fibrillation is a kind of cardiac arrhythmia in which the electrical signals of the heart are uncoordinated. The prevalence of this disease is increasing globally and the curative treatment for this problem is catheter ablation therapy. The adequate contact force between the tip of a catheter and cardiac tissue significantly can increase the efficiency and sustainability of the mentioned treatment. To satisfy the need of cardiologists for haptic feedback during the surgery and increase the efficacy of ablation therapy, in this paper a sensor-free method is proposed in such a way that the system is able to estimate the force directly from image data. To this end, a mechanical setup is designed and implemented to imitate the real ablation procedure. A novel vision-based feature extraction algorithm is also proposed to obtain catheter's bending variations obtained from the setup. Using the extracted feature, machine learning algorithms are responsible of estimating the forces. The results revealed $MAE < 0.0041$ and the proposed system is able to estimate the force precisely.

***Index Terms*—** Cardiac catheter, machine learning, regression, artificial neural network, force estimation, machine vision.

1. INTRODUCTION

Cardiovascular Diseases (CVDs) as one of the main reasons of global mortality are caused by disorders of the cardiovascular system [1]. According to heart disease statistics, the annual worldwide deaths associated with CVDs are over 17 million [2]. Atrial Fibrillation (AF) is a common heart arrhythmia which occurs due to erratic electrical impulse of the heart and has affected at least 3 to 6 million people in the United States [3].

Catheter ablation is a well-known minimally invasive treatment for AF to locally heat and destroy (ablate) arrhythmogenic cardiac tissue [4, 5, 6]. To perform the ablation treatment, a long flexible tube called catheter, is inserted into the vascular system to deliver some source of energy to the arrhythmia spots of the heart under X-ray fluoroscopy or MRI monitoring [7].

Adequate catheter-tissue Contact Force (CF) is known as a procedural success factor that leads to a sustainable effect of catheter therapy [8]. Accordingly, the force sensing system plays a significant role in cardiac catheterization [9]. In accordance with experimental studies, CF between $0.1N$ and $0.3N$ is a safe and effective range [10]. Besides, image-based position tracking of the catheter shaft and tip is considered as an important feature in terms of accuracy of guidance [11].

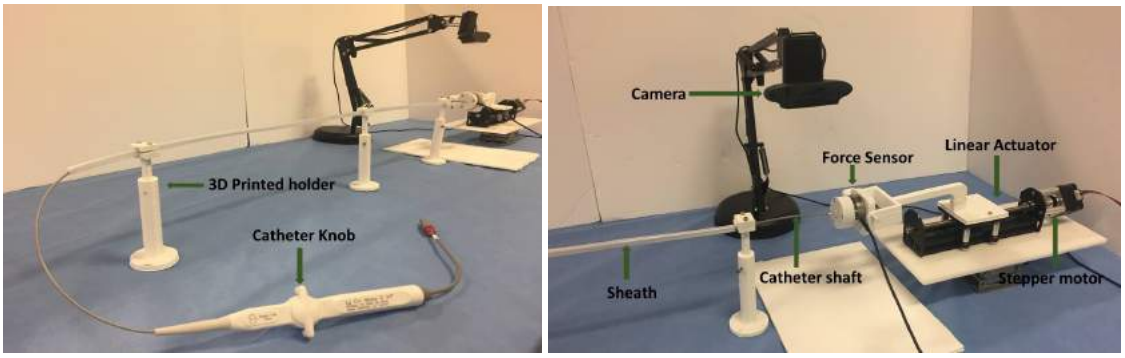
Sensor-based and sensor-free approaches are proposed methods for measuring the CF of a catheter's distal tip [12, 13]. Despite the fact that sensor-based methods are providing accurate measurement, implementation of tactile sensors in catheters has some challenges including high-end cost, physical issues ,and also difficulties in data acquisition systems in the unstructured environment [14]. Accordingly, as an alternative, sensor-free methods have caught attentions in the literature [12].

One approach of the sensor-free method is the analysis of catheter shape in which a parameter called "force index" is identified to address the force range [15]. However, based on experimental results, this method is not capable of detecting the full range of forces. Another approach is model-based techniques consisting of beam theory models, Cosserat-type rod theories ,and multi-body dynamics [16, 17, 18]. Although the mentioned model-based manners in some cases provide an accurate estimation of the force, the main effective factor for this accuracy is the optimal model parameters [19]. In another study, Runge *et al.* [20], used a finite element model to train an artificial neural network for soft robotic application. However, this method requires accurate finite element modeling in which the material parameters and manufacturing aspects of the soft robot should be considered carefully.

Overall, the accuracy of the model-based methods highly depends on the model parameters. In addition, they are computationally expensive, especially when it comes to detection of a real catheter from the operating room's monitor as these models need information from the image of the catheter.

In this study, we proposed a new vision-based solution in collaboration with machine learning methods to address the CF issue in ablation catheters. This system is functional as

This research was supported by the Natural Science and Engineering Research Council (NSERC) of Canada through CREATE Grant for Innovation-at-the-Cutting-Edge(ICE) and Concordia University, Montreal, QC, Canada.



(a) Overall view of setup (b) A view of Catheter, Sensor, Actuator and Camera

Fig. 1: Experimental Setup

a sensor-free and real-time method in which the CF can be estimated directly from the image data. As the deflectable distal shaft of the catheter has unique bending under applied forces, this behaviour can be considered as a distinguishable factor of force estimation [15]. Accordingly, in this work, a mechanical setup has been designed and implemented in order to simulate the authentic Operation Room (OR) along with the available tools for performing an ablation task. Using the data captured from the aforementioned setup, a machine vision algorithm is devised as a feature extractor to find the points on the catheter's deflectable distal shaft within images. These points are deemed as the features which translate the catheter's tip into a numerical feature space. Subsequently, multiple architectures for Artificial Neural Networks (ANN) have been designed and implemented so as to model the extracted features and map every catheter's image to its corresponding contact force. In addition to ANNs, we model the data using Support Vector Regression (SVR) with the aim of making a benchmark. These models are considered as a system which maps the features to the CF in x and y direction. In the next section, the developed experimental setup and data compilation will be explained. Then, the methodology including the feature extraction algorithm as well as the modeling methods will be elaborated. The paper will be concluded in the last section.

2. EXPERIMENTAL SETUP AND PROCEDURE

The experimental setup used for data collection is shown in Fig. 1. In this setup, the camera plays the role of X-ray fluoroscopy machine in a real OR. In addition, a motorized linear actuator simulates the heart motion (one direction) in which an attached force sensor is recording the CF.

2.1. Experimental setup design

The experimental setup is designed in an attempt to collect images from the deflectable shaft of the catheter and corresponding forces. Fig. 1 presents the setup consisting of a 1-DOF (Degree Of Freedom) linear actuator equipped with a 2-phase stepper motor (17HS4401-S 40mm Nema) powered by a micro-step driver (HANPOSE TB660), a Camera (C920 for 640×480 pixels resolution), a 6-DOF Force sensor (ATI Mini40), a Bi-directional catheter (Boston Scientific Blazer II XP), 3D printed parts for holders and a plastic sheath. In this setup, the catheter is passed through a plastic sheath in a straight path in which the sheath is fixed by 3 holders and the Knob of the catheter is configured at zero degrees. The deflectable section at its base point where the body of the catheter is connected to the bending section is fixed by the 3rd holder. Hence, it cannot move inside the sheath. The force sensor attached to the 1-DOF motorized actuator is used to measure the applied forces at the tip of the catheter. In addition, the camera is implemented perpendicular to the bending section to capture a planar image for every sample.

2.2. Data Collection

Fig. 2 depicts the interaction between the software and the hardware in the experimental setup to collect a dataset comprising of 2000 sample images from deflected shaft under the applied forces by the actuator. After calibration of the sensor, data collection is done by following steps for each sample:

1. A Computer program developed in Python sends a command to an Arduino UNO to manipulate the motorized linear actuator.
2. The Arduino and stepper motor driver control the stepper motor rotation for three micro-steps equivalent to 0.6 degree (the driver is set to 1600 pulse per revolution to reach 0.2 degree per pulse).
3. Afterward, the Arduino sends an acknowledge signal to the computer.
4. The image of the deflected shaft of the catheter is captured.

5. The force data (two directions) is recorded from a 16-bit data acquisition device (USB 6210, National Instruments, Austin, TX).

This procedure is repeated 2000 times to build a dataset that contains deflected shaft images from the initial shape of the bendable shaft to fully deflected formation and their corresponding forces.

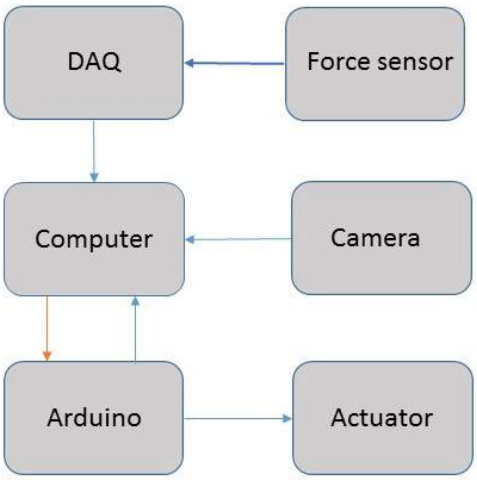


Fig. 2: Software and hardware interaction

3. METHODOLOGY

3.1. Feature extraction

Monitoring the catheter's deflection (distal shaft) during the ablation process provides information about CF [15]. In this regard, a real-time machine vision algorithm is required to extract features and translate the deflection in a numerical feature space. In this study, we have proposed an algorithm to obtain ordered points as features on the body of the catheter's deflectable section. The **Algorithm 1** shows the image processing procedure to extract ordered points from the base point of the deflectable shaft to the tip of catheter. An image captured by the camera is presented in Fig. (3a). This image is converted to the gray scale format and then a binary threshold operation is applied to segment the catheter in the image [see Fig. (3b)]. Subsequently, the algorithm vertically searches through the matrix of the 2D binary image until it finds the body of the catheter. At this point, the Cartesian coordinates (x and y) are stored as the first feature. This procedure continues until the last possible vertical search path is met. The distance between each vertical search path is a hyper-parameter that defined as the skip point. This criteria denotes the number of features. For instance, every image is represented in a 106-dimension feature space, if the skip point is equal to 5. Fig. (3c) depicts the overall procedure of vertical search where blue lines indicate the search path. Fig.(3d) shows the extracted ordered points (features) on the catheter in which the tip as the last recorded point is detected.

The aforementioned algorithm is applied to 2000 images and multiple datasets have been generated with different values for the skip point. Using the compiled datasets, the machine learning models are responsible of estimating the forces.

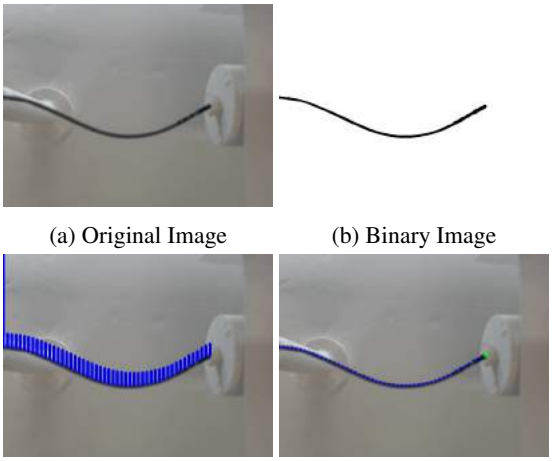
Algorithm 1: Feature Extraction Algorithm

```
Input:RGB image from the Camera
Output:Ordered points on the body of catheter
GrayscaleFunction(RGBimage)
ThresholdFunction(Grayimage)
Flag =False
j=0
for i = 0; i <= width; i = i + 5 do
    for j; j <= height; j+ = 1 do
        Flag =False
        pixel location=[i,j]
        if pixel value ==0 then
            Record i,j location of the point
            Flag=True
            j=horizontal location of the point+30
            Break the Vertical(height)search loop
        else
            Continue the Vertical(height) search
        end
    end
    if Flag==False then
        Break the Horizontal(width) step loop
    else
        Continue the Horizontal(width) search
    end
end
```

3.2. Machine Learning Models

Having the output data of the feature extraction phase, a modeling method is required in between so as to map every single record of the dataset to its corresponding force value. Since the goal is to estimate the forces directly from the images, the modeling technique deals with continuous values. In contrast to the modeling of quantitative values as a classification problem, in the current work, the system strives to create a regression over the data. With this in mind, two more popular modeling approaches have been intended with multiple configurations in order to approximate the forces in x and y direction: Artificial Neural Network (ANN) and Support Vector Regression (SVR) [21, 22].

The ANN is considered to receive a feature vector $x \in$



(a) Original Image (b) Binary Image

(c) Search Path (d) Points

Fig. 3: Features Extraction

R^m where m is the number of features depending on the the settings of the feature extraction phase. It is also expected to output the result in a two-dimensional space in order to estimate the actual forces in x and y direction. To this end, a combination of the following stacked layers is deemed so as to design the architecture of the ANN. The input feature vector x goes into a dense layer:

$$n_l = \sigma(W_l n_{l-1} + b_l) \tag{1}$$

Where W and b are the weights and biases (neurons) respectively. l denotes the layer of the network and n_l is the output of the current layer l. In addition, n_{l-1} indicates the output of the previous layer while for the first layer $n_{l-1} = x$. Also, σ is the activation function in such a way that for the intended architecture the $ReLU$ is preferred:

$$\sigma(z) = max(0, z) \tag{2}$$

Obviously, the activation function above squeezes the negative value and replace them with 0. Moreover, it is common in the Deep Learning (DL) and ANN to equip the model with the regularization term, e.g., "L1 or L2-Regularization" with the aim of preventing the model from over-fitting. Here in this design, the Dropout is opted as the alternative to the regularization method above. In this method, a coefficient δ is multiplied by every component of the W and b matrices and calculated as follows [23]:

$$\delta = \begin{cases} w_j, b_j & \text{with P(r),} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

In fact, the component j of the parameter matrix W or b is remained with the probability $P(r)$ where $P \backsim Bernoulli(r)$ and in this architecture $r = 0.1$. It is worth noting that, Batch Normalization (BN) method is widely used in ANN in order to accelerate the optimization process [24]. In this work, the design incorporates the BN to the graph prior to applying the activation function. The last layer of the network outputs a vector with two elements corresponding to x and y forces. To optimize the model and train the neurons, Mean Absolute Error (MAE) is selected as the loss function of the model:

$$MAE(n_l, f) = \frac{\sum_{i=1}^{n} |n_{l,i} - f_i|}{n} \tag{4}$$

In the equation above, f denotes the actual forces that the model is supposed to estimate. The designed network can be optimized using Adam optimizer [25].

For the sake of benchmarking the eclectic number of modeling methods for the regression problem upon the extracted features, a linear Support Vector Regression (SVR) has been utilized in order to generate the forces out of the the given features [22]. Since the SVR is a well-known traditional Machine Learning (ML) algorithm for which it has widely contributed to a broad range of applications and also there are valuable references in the literature about it, in this work, we will not explain the method in details. Two separate SVR models are considered for this problem: the SVR for the force in x direction and the other one for y direction. It is worth to say that, the ANN and all derived configurations have been implemented on Python using Tensorflow 2.4 while the implemented API of Sklearn has been utlized for the SVR models [26, 27]. In contrast to the ANN models, no further modification or implementations have been done for the SVR models.

4. RESULTS AND DISCUSSION

In this section, the performance of the proposed system has been investigated in both the feature extraction and the modelling phase. To this end, a diverse range of configurations has been designed and the results have been compared using three main metrics: MAE, MSE and R2. The feature extraction module was fed by a dataset containing 2000 images of the catheter's tip. The dataset obtained from the feature extraction phase was normalized and divided into three sub-datasets: a training set encompassing 1280 samples, a testing set containing 400 samples and a validation including 320 samples. As reported in Table 1, 10 different configurations have been design so as to compare the performance of the proposed method accurately. For all NN-based methods every batch of the dataset includes 16 samples and the system was trained in 200 epochs while the learning rate $lr = 0.001$. The first configuration is a graph of 9 dense layers as a feed-forward NN in which the layers contains the following neurons respectively: 256, 256, 128, 128, 64, 64, 32, 32, 2. This model trained on the dataset comprising 36 features acquired from the feature extraction algorithm with 15 skip points. The

Table 1: This table compares the performance of multiple modeling methods with different configurations.

No.	method	configuration	layers	feature dim	skip points	*MAE*	*MSE*	*R2*
1	DNN	dense	9 layers	36	15	0.0042	3.4868e-05	0.984
2	ANN	dense	[128, 64, 2]	36	15	0.0041	3.5274e-05	0.986
3	ANN	dense	[128, 64, 2]	106	5	0.0040	4.4928e-05	0.978
4	ANN	dense	[128, 64, 2]	22	24	0.0043	3.0972e-05	0.986
5	ANN	dense + dropout	[128, 64, 2]	36	15	0.0606	0.00515	0.218
6	ANN	dense + batch	[128, 64, 2]	36	15	0.0044	5.0931e-05	0.974
7	ANN	dense + batch + dropout	[128, 64, 2]	36	15	0.0392	0.0022	0.328
8	SVR	linear	-	22	24	0.0068	3.0705e-04	0.913
9	SVR	linear	-	36	15	0.0058	1.9133e-04	0.9449
10	SVR	linear	-	106	5	0.0046	6.2738e-05	0.982

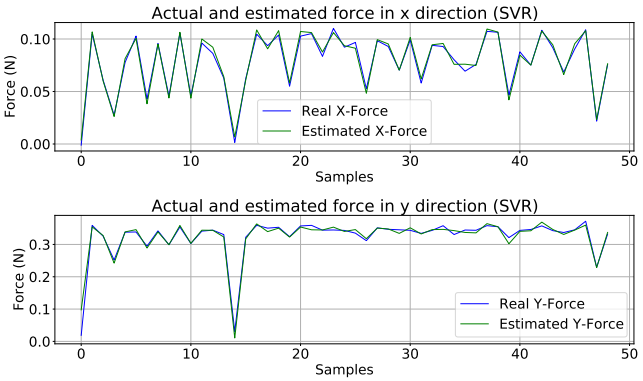


Fig. 4: The performance of the SVR for approximating the forces in both x and y direction.

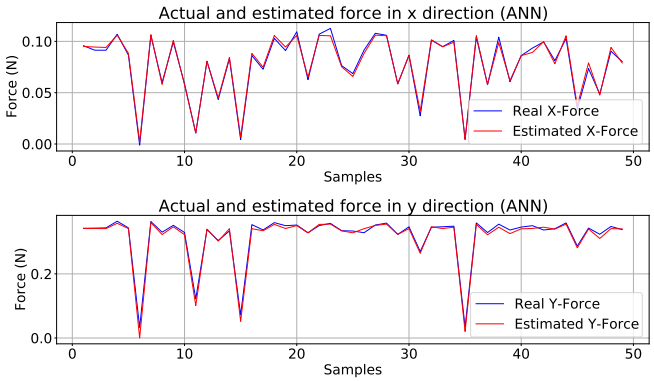


Fig. 5: This plot shows the performance of the designed ANN in estimating the forces in x and y direction.

second configuration has a shallow architecture while the layers are analogous to the previous Deep NN in terms of layer's type. This NN showed a better accuracy whereas the parameters was considerably decreased. One reason is that the complexity of the captured data was not significant so a shallow NN was capable of extracting the model properly. However, the deeper network needs more epochs to be trained completely. Configuration 2 to 4 investigated the impact of feature extraction phase and the feature dimension in the modeling. As it can be seen, given the fact that the training epochs was the same for all configurations, the representation of the catheter images in higher dimensions did not reached to an acceptable performance. However, the data with more features not only needs more parameter and epochs to be trained but also it requires a deeper NN to obtain the relationship between every dimension. Configuration 5 to 7 inspected the influence of adding batch normalization and dropout to the network's graph. The dropout layer prevent the model from reaching to the convergence point. It is worth noting that, the training loss was tracked on the validation set during the training phase and no evidence of over-fitting was caught.

The last three configurations in Table 1 reported the performance of a linear SVR model on the datasets. For every dataset, two separate SVR models were trained: one for the

estimation of forces in x direction and the other one for y direction. The reported results are the average of corresponding metrics for x and y models. The tolerance for stopping the training procedure was set to 0.001. Given the tolerance value, the system keeps learning until the tolerance becomes satisfied. For this reason, the impact of representing data in different dimension size can be tangibly evaluated. The SVR showed a better regression's performance on the dataset with higher dimensions while the over performance of the ANN methods surpassed. Fig. 4 and 5 demonstrate the performance of configuration 2 and 10 respectively and show the proposed system can estimate the actual forces accurately.

5. CONCLUSIONS

In this work, a sensor-free method has been proposed for estimating the force of the catheters' tip with the aim of contributing to the catheter ablation treatment for cardiovascular diseases. The method is capable of approximating the forces directly from the images. To this end, a mechanical setup has been designed and implemented in order to imitate an authentic operation room for catheter ablation. Using the setup, the system compiled a dataset containing the images of a catheter's tip and the forces associated to every image. A novel feature extraction algorithm has been proposed to extract the variation of catheter's deflection within the images and represent them in the multi-dimensional feature space. Having the dataset of extracted features, different feed-forward neural network has been designed and implemented to make a regression over the data. Besides, Support Vector Regression as a conventional machine learning method was deployed to model the data as well. The output of the proposed feature extraction collaborating with the implemented modeling methods estimated the forces precisely. As the future work, we will extend the current system in such a way that the 3D forces can be estimated directly from the images without a feature extraction phase.

6. REFERENCES

[1] World Health Organization website, "Cardiovascular diseases," .

[2] Salim S Virani, Alvaro Alonso, Emelia J Benjamin, Marcio S Bittencourt, Clifton W Callaway, April P Carson, Alanna M Chamberlain, Alexander R Chang, Susan Cheng, Francesca N Delling, et al., "Heart disease and stroke statistics—2020 update: a report from the american heart association," *Circulation*, vol. 141, no. 9, pp. e139–e596, 2020.

[3] Jelena Kornej, Christin S Börschel, Emelia J Benjamin, and Renate B Schnabel, "Epidemiology of atrial fibrillation in the 21st century: novel methods and new insights," *Circulation Research*, vol. 127, no. 1, pp. 4–20, 2020.

[4] Pablo B Nery, Rebecca Thornhill, Girish M Nair, Elena Pena, and Calum J Redpath, "Scar-based catheter ablation for persistent atrial fibrillation," *Current opinion in cardiology*, vol. 32, no. 1, pp. 1–9, 2017.

[5] Gjin Ndrepepa and Heidi Estner, "Ablation of cardiac arrhythmias—energy sources and mechanisms of lesion formation," in *Catheter Ablation of Cardiac Arrhythmias*, pp. 35–53. Springer, 2006.

[6] Dieter Haemmerich, "Biophysics of radiofrequency ablation," *Critical Reviews™ in Biomedical Engineering*, vol. 38, no. 1, 2010.

[7] Xiaohua Hu, Ang Chen, Yigang Luo, Chris Zhang, and Edwin Zhang, "Steerable catheters for minimally invasive surgery: a review and future directions," *Computer Assisted Surgery*, vol. 23, no. 1, pp. 21–41, 2018.

[8] Luigi Di Biase, Andrea Natale, Conor Barrett, Carmela Tan, Claude S Elayi, Chi Keong Ching, Paul Wang, AMIN AL-AHMAD, Mauricio Arruda, J David Burkhardt, et al., "Relationship between catheter forces, lesion characteristics,"popping," and char formation: experience with robotic navigation system," *Journal of cardiovascular electrophysiology*, vol. 20, no. 4, pp. 436–440, 2009.

[9] Dipen C Shah and Mehdi Namdar, "Real-time contact force measurement: a key parameter for controlling lesion creation with radiofrequency energy," *Circulation: Arrhythmia and Electrophysiology*, vol. 8, no. 3, pp. 713–721, 2015.

[10] Nilshan Ariyarathna, Saurabh Kumar, Stuart P Thomas, William G Stevenson, and Gregory F Michaud, "Role of contact force sensing in catheter ablation of cardiac arrhythmias: evolution or history repeating itself?," *JACC: Clinical Electrophysiology*, vol. 4, no. 6, pp. 707–723, 2018.

[11] Stijn De Buck, Joris Ector, Andre La Gerche, Frederik Maes, and Hein Heidbuchel, "Toward image-based catheter tip tracking for treatment of atrial fibrillation," in *CI2BM09-MICCAI Workshop on Cardiovascular Interventional Imaging and Biophysical Modelling*, 2009, pp. 8–pages.

[12] Kai Xu and Nabil Simaan, "An investigation of the intrinsic force sensing capabilities of continuum robots," *IEEE Transactions on Robotics*, vol. 24, no. 3, pp. 576–587, 2008.

[13] PJ French, D Tanase, and JFL Goosen, "Sensors for catheter applications," *Sensors Update*, vol. 13, no. 1, pp. 107–153, 2003.

[14] Liang Zou, Chang Ge, Z Jane Wang, Edmond Cretu, and Xiaoou Li, "Novel tactile sensor technology and smart tactile sensing systems: A review," *Sensors*, vol. 17, no. 11, pp. 2653, 2017.

[15] Mahta Khoshnam and Rajni V Patel, "Estimating contact force for steerable ablation catheters based on shape analysis," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 3509–3514.

[16] Mahta Khoshnam, Mahdi Azizian, and Rajni V Patel, "Modeling of a steerable catheter based on beam theory," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 4681–4686.

[17] John Till, Vincent Aloi, and Caleb Rucker, "Real-time dynamics of soft and continuum robots based on cosserat rod models," *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 723–746, 2019.

[18] Robert J Webster III, Joseph M Romano, and Noah J Cowan, "Mechanics of precurved-tube continuum robots," *IEEE Transactions on Robotics*, vol. 25, no. 1, pp. 67–78, 2008.

[19] Shahir Hasanzadeh and Farrokh Janabi-Sharifi, "Model-based force estimation for intracardiac catheters," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 1, pp. 154–162, 2015.

[20] Gundula Runge, Mats Wiese, and Annika Raatz, "Fem-based training of artificial neural networks for modular soft robots," in *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2017, pp. 385–392.

[21] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, MIT Press, 2016, http://www.deeplearningbook.org.

[22] Mariette Awad and Rahul Khanna, *Support Vector Regression*, pp. 67–80, Apress, Berkeley, CA, 2015.

[23] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014.

[24] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015.

[25] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," 2017.

[26] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, Software available from tensorflow.org.

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

A CLASSICAL MACHINE LEARNING APPROACH FOR EMG-BASED LOWER LIMB INTENTION DETECTION FOR HUMAN-ROBOT INTERACTION SYSTEMS

Hasti Khiabani, Mojtaba Ahmadi

Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON, Canada

ABSTRACT

Surface Electromyography (sEMG)-based intention-detection systems of lower limb can intelligently augment human-robot interaction (HRI) systems to detect subject's walking direction prior-to or during walking. Ten Subject-Exclusive (Subj-Ex) and Generalized (Gen) Classical Machine Learning (C-ML)-based models are employed to detect direction intentions and evaluate inter-subject robustness in one knee/foot-gesture and three walking-related scenarios. In each, sEMG signals are collected from eight muscles of nine subjects during at least nine distinct gestures/activities. Linear Discriminant Analysis (LDA) and Random Forest (RF) classifiers, applied to the Time-Domain (TD) feature set (of the four input sets), provided the best accuracy. Subj-Ex approach achieves the highest prediction accuracy, facing occasional competition from the Gen approach. In knee/foot gesture scenario, LDA reaches an accuracy of 91.67%, signifying its applicability to robotic-assisted walking, prosthetics, and orthotics. The overall prediction accuracy among walking-related scenarios, though not as remarkably high as in the knee/foot gesture recognition scenario, can reach up to 75%.

Index Terms— surface EMG - Intention Detection - Direction Detection - Classical Machine Learning - intelligent HRI- Robotic-Assisted Walking

1. INTRODUCTION

People with mobility-related disabilities including Cerebral Palsy (CP) patients can observe improvements in their quality of life through incorporation of various types of robotic devices, e.g. rehabilitation, assistive, and human computer interaction (HCI) devices [1]. To be compliant to users' needs, such robots should be intelligent enough to detect their human partners' intention to be able to provide beneficial and engaging feedback to them in executing desired motion [2]. The study of human pattern recognition based on C-ML methods with a classification approach roots back to 1993 with Hudgin's study on classifying four hand gestures from two sEMG channels with Multi-layer Perceptron (MLP) classifier based on five TD features [3]. Researchers then investigated other classifiers with various feature sets to improve the classification. Englehart et al. stated that the performance of feature extraction and dimensionality reduction is highly dependent on the structure of a classifier [4]. Later, they demonstrated that four channels of EMG data considerably enhance the classification accuracy compared to one or two channels [5]. Further, they reduced the error in prediction by applying a post-processing technique, called Majority Vote, to reduce the number of false predictions [6]. The first study that successfully achieved a 70-80% classification accuracy on a large number of hand gestures (52 gestures on the Ninapro dataset) was proposed by Kuzborskij et al. They used an SVM classifier on a set of TD and Frequency-Domain (FD) features extracted from 8 channel of myoelectric signals [7, 8]. Atzori et al. improved upon the previous research by using an ensemble RF classifier to achieve an average accuracy of 75.32% on a linear combination of features [9]. Gijsberts improved the performance of classification by 5% by combining extracted features from sEMG and acceleration signals [10].

There is a broad and growing body of literature on lower limb gesture recognition and intention detection to control assistive robots using sEMG signals [11, 12, 13, 14]. Lyons et al. illustrated that more than 90% accuracy across seven gestures can be achieved for foot gesture recognition [15]. They considered an isometric contraction in the patterns of the lower limb, while most other studies on lower leg intention detection deal with isotonic contractions in walking practices [14], e.g. gait phase recognition and mode predictions(sitting, standing, ramp ascent/descent, etc.)[16, 13]. Researchers also investigated the effects of sensor choice, fusion, and configuration (unilateral or bilateral). EEG sensors were utilized to detect a healthy subject's intentions to perform six movement tasks for eventual combination with exoskeletons for neurorehabilitation [17]. Zhang et al. fused mechanical measurement data from Ground Reaction Force (GRF) and Kinematic information and neuromuscular data to identify three activities (level-ground walking, sitting, and standing) in a study on patients with multiple sclerosis. Their designed system was capable of reliably predicting the users' intention in static states while also being able to correctly predict the activity transitions about 100 to 130 ms before the actual transition [13]. Hu et al. used bilaterally collected EMG data and kinematics

This research was supported by the Discovery Grant, from Natural Sciences and Engineering Council Canada (NSERC) held by Mojtaba Ahmadi. Support for Hasti Khiabani was provided by the NSERC CREATE Grant 497303.

of joints and limbs to detect the smooth transition between locomotor activities and demonstrated that fusing different sensors while taking into account the manner in which bilateral sensor data is collected can enhance the classifiers' ability to predict the true activity [18]. This study extends previous lower limb intention detection studies by solely relying on sEMG data to detect a more comprehensive range of gestures/activities that correspond to specific directions in different phases of movement (prior-to or during walking) by exploring scenarios that cover a broad range of motion patterns. C-ML is utilized as a preliminary method to identify users' intended direction and inter-subject robustness of the models are evaluated by employing two implementation strategies (Gen and Subj-Ex).

2. EXPERIMENT

A set of experiments has been designed to collect eight channels of sEMG data with sample rate of 1kHz through dry electrodes using Delsys Bagnoli system. Sensors where mounted by the structure proposed in [19] across the Thigh and Shank of the dominant leg. Data is collected from nine subjects in five repeating cycles. Each cycle contains at least five seconds of data from the subject performing one randomly selected gesture in each scenario. The research was cleared under ID: 114417, by Carleton University's Research Ethics Board and the informed consent was obtained from all subjects.

2.1. Scenarios

Early into the data collection process, it was observed that there is a distinction between users' motion patterns in (i) the initial intention phase, (ii) the initiation of motion phase, and (iii) the phase where the user is in motion. The more explicit the distinction between these phases in the training is, the more informative the model will become. Thus, three walking scenarios plus a knee/foot-gesture based verification scenario are planned to investigate each of the phases. The verification scenario involves isometric knee/foot gestures in static position that closely resemble the hand/arm gestures commonly evaluated in the existing literature [7, 9]. As so, notable performance of the model in this scenario is a promising sign of its applicability to more complex gestures/activities. Scenarios are detailed below and illustrated in Figs. 1, 2, and 3. In the three walking scenarios each experimented gesture corresponds to a movement in one of the 9 classified directions.

- **Standing Position Scenario (StPS)** comprises of the subject taking a step towards the desired direction and holding that gesture (shown by a GUI system). Appropriately, this scenario is categorized as a static scenario.
- **Attempted Motion Scenario (AMS)** the subject imagines and generates the forces necessary to initiate the motion towards the desired direction without actually moving. This scenario is also static due to the absence of motion.
- **Dynamic Motion Scenario (DMS)** where the subject goes through the entire motion that involves the subject intending, initiating and walking with their average speed towards the desired direction. This scenario is categorized as a dynamic scenario.
- **Sitting Position Scenario (SiPS)** which was foreseen to facilitate the verification of the models by having the subject perform some static isometric gestures while sitting.

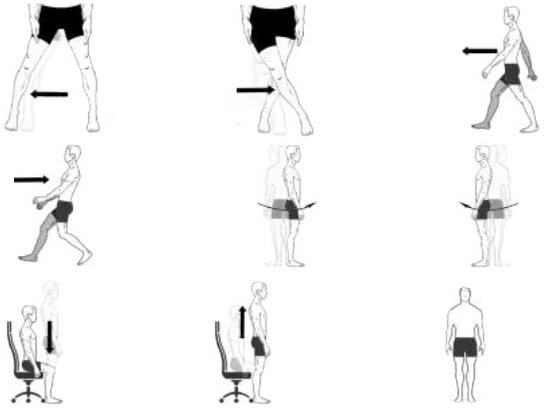


Fig. 1: Gestures/activities performed in DMS and AMS consisting of: side step, cross step, forward, backward, internal rotation, external rotation, sit, stand, and rest.

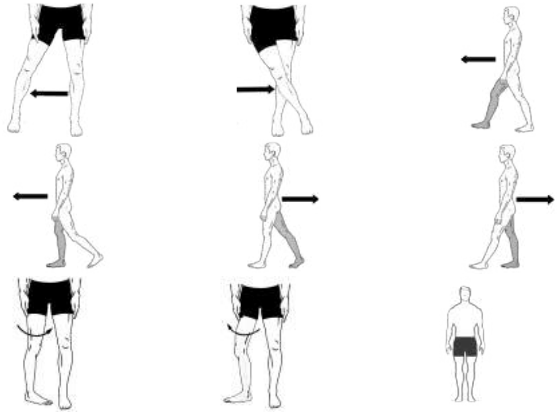


Fig. 2: StPS gestures involving: side step, cross step, forward, forward2, backward, backward2, hip internal rotation, hip external rotation, and rest.

3. FEATURE EXTRACTION AND LEARNING METHODOLOGY

The collected data is segmented into windows of 260 ms with 235 ms overlap and filtered by a butterworth band-pass filter of 20-450Hz. Feature engineering has been chosen as the feature extraction method. In general, TD features have shown better performance in comparison with FD features for classification tasks with less required computational power causing less delay on real-time applications [20]. The most commonly used features for avoiding redundancy while preserving accuracy for pattern recognition tasks are considered:

Fig. 3: SiPS gestures involving: ankle dorsiflexion, plantarflexion, inversion, eversion, lateral rotation, and medial rotation, knee abduction, adduction, flexion, extension, and rest.

Mean Absolute Value (MAV), Zero Crossing (ZC), Waveform Length (WL), Slope Sign Change (SSC), Root Mean Square (RMS), integrated Absolute Value (iEMG), and Skewness (Skew). Further, Auto-Regressive coefficients (AR) and Histogram (HIST) features that can convey FD information of the signals are also added to investigate their effectiveness in each group. Principle Component Analysis (PCA) is also applied to investigate its capability to chose the most important features and neglect less important ones (20% of features are neglected). Table 1 summarizes different combinations of features used as inputs of C-ML models.

Table 1: List of features extracted to be fed to C-ML

List of Manually Selected Features		
Abbreviation	Features	PCA
TD	MAV,SSC,ZC,WL,RMS,AR	No
TD-PCA	MAV,SSC,ZC,WL,RMS,AR	Yes
Enhanced TD	MAV,SSC,ZC,WL,RMS,AR,iEMG,HIST,SKEW	No
Enhanced TD-PCA	MAV,SSC,ZC,WL,RMS,AR,iEMG,HIST,SKEW	Yes

C-ML contains a wide family of classifiers that are mostly built upon statistics and probabilistic reasoning. Ten of these classifiers are chosen to be investigated for intended direction detection involving: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), LDA and Quadratic Discriminant Analysis (QDA), Decision Trees(DT), Gaussian Naive Bayes(GNB), Ensemble Bagging models (consisting of RF and Bag), and ensemble Boosting (consisting of Adaptive Boosting(Ada) and Gradient Boosted DT(GBDT)).

4. IMPLEMENTATION STRATEGIES

In order to generalize the classifiers and ensure their prediction robustness against the choice of the subject, two implementation strategies are designed. Both strategies single out the same cycle for eventual testing (out of total of five cycles), but they train their models differently.

- **Generalized (Gen):** Out of the four remaining cycles, three are concatenated to form the training set and the fourth cycle is set aside as validation set. The classifier is optimized through cross validation over the multiple choices of the validation set out of these four cycles. The reported accuracy is averaged over these iterations.
- **Subject-Exclusive (Subj-Ex):** Each classifier is trained and validated per each subject separately in the same manner described in Gen. The accuracy of the trained model is then evaluated on the test set of that same subject. The overall performance of the model is reported by averaging these individual accuracies.

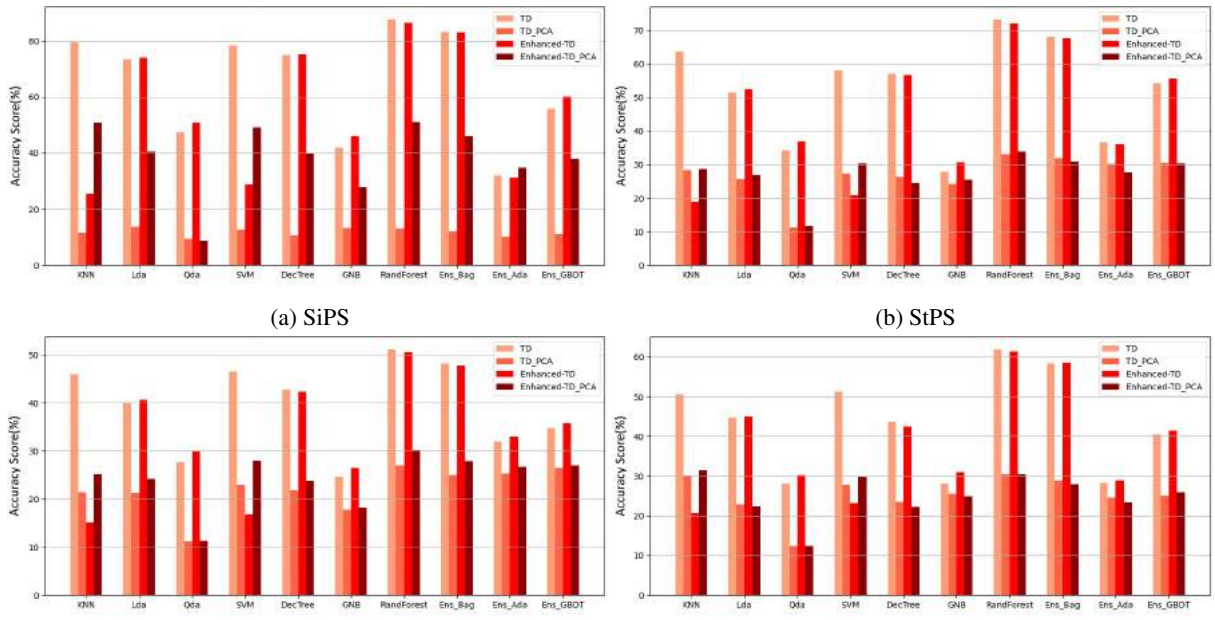
5. RESULTS

Ten C-ML models with two different implementation strategies are applied on each of the computed four feature sets and the model accuracy is reported among these 80 cases of each of the four scenarios. Fig 4. depict an analysis of different feature sets in each scenario found by Gen strategy, compared in terms of: (i) classifier, (ii) whether PCA is applied or not, and (iii) experiment scenarios. Table 2, depict a comparison of two different strategies that are used to investigate the generalization capabilities of each model. The TD feature set is used to carry out this analysis because of its superior performance in all scenarios as can be seen in Fig. 4. Because of the almost balanced distribution of the data sets, the chance of a random guess being correct is 9.1% and 11.11% in SiPS and three walking-related scenarios, respectively% .

Table 2: Model's Accuracy Score on test set in each scenario with TD features

	SiPS		**StPS**	
Classifier	Gen(%)	Subj-Ex(%)	Gen(%)	Subj-Ex(%)
KNN	79.43	83.63	63.72	67.06
LDA	73.34	91.67	51.48	75
QDA	47.4	82.98	34.17	64.91
SVM	78.43	85.51	58.08	68.29
DT	74.92	79.94	57.14	63.08
GNB	42.03	82.1	27.74	62.84
RF	87.63	87.2	73.32	73.38
Bag	83.12	83.81	68.1	68.35
Ada	31.98	27.46	36.69	29.1
GBDT	55.76	68.64	54.28	60.26
	AMS		**DMS**	
Classifier	Gen(%)	Subj-Ex(%)	Gen(%)	Subj-Ex(%)
KNN	46.04	48.90	50.59	55.06
LDA	39.87	53.1	44.73	61.82
QDA	27.67	49.56	28.06	45.54
SVM	46.49	53.43	51.17	58.9
DT	42.73	47.13	43.62	51.12
GNB	24.65	50.61	28.02	39.62
RF	51.13	53.81	61.89	64.8
Bag	48.15	51.52	58.34	59.76
Ada	31.88	31.1	28.19	25.38
GBDT	34.79	45.52	40.58	49.86

In the **SiPS** with 11 distinct gestures, The LDA (91.67%) and the RF (87.63%) classifiers achieved the highest accuracy among Subj-Ex and Gen methods, respectively. Due to



(a) SiPS (b) StPS

(c) AMS (d) DMS

Fig. 4: Each classifier's performance when different set of features are selected in each scenario when Gen strategy is employed.

the high accuracy obtained in this scenario, they can be reliably used in various applications. In the **StPS** with 9 distinct gestures, LDA (75%) and RF (73.32%) classifiers achieved the highest accuracy among Subj-Ex and Gen strategies, respectively. In the **AMS** with 9 distinct gestures, RF achieved the highest accuracy among Subj-Ex strategies with 53.81% and Gen strategies with 51.13%, respectively. Analyzing each classifier's results when trained in each mode shows us that there is a noticeable difference in their ability to differentiate between various gestures. The cause can be sought out by taking a close look at the experiment itself. During the experiment, the subjects were asked to imagine and provide the muscle tension required as if they were supposed to move in a specific direction without actually performing the task. Each person has a distinct manner of providing such an intention while abstaining from motion. Thus, an explicit difference in the sEMG signal between different subjects is expected. This best resembles the behavior of a subject with some motor-disabilities. Thus, in such a case, Subj-Ex methodology is the most advantageous to enhance the detector's accuracy level. In the **DMS** with 9 distinct activities, RF surpassed all other classifiers either trained with Subj-Ex, or with Gen strategy with the accuracy of 64.8%, and 61.89%, respectively. Lower accuracy was expected in this scenario due to the activities' more complex dynamic nature. The inherent dynamic complexities make it difficult for C-ML classifiers to relate different data windows and differentiate between correlated tasks. Overall, RF, LDA, Bag, and SVM were the top performers across all scenarios. Nevertheless, RF has been the best classifier for almost all of our scenarios. Furthermore, the number of gestures to be identified in the StPS, AMS, and DMS scenarios was lower than SiPS. Still, the classifiers were not as highly capable of distinguishing between different types of gestures. The cause of this can be rooted back to the muscles recruited to record sEMG data. The most involved muscles for SiPS gestures are the ones that are being monitored. On the other hand, in other scenarios, many gestures originate from the hip muscles. Due to the extra burden imposed on subjects to attach hip muscles, these muscles were not monitored during the experiment. Lack of information from hip muscles weakens the classifier in distinguishing the tasks that are highly dependant on it.

6. CONCLUSION

Overall, SiPS models are demonstrated to reliably facilitate knee/foot-related intelligent assistive devices and joystick-type controllers, especially if the subjects are trained to use their foot gesture as controller more frequently. AMS-related scenarios are better when trained on each subject exclusively. The study presented a promising early proof of concept. Although a very high detection accuracy is not achieved in StPS, AMS and DMS, the detection model can be further adjusted to be properly integrated into the control systems of robotic-assisted walking devices by using additional processing methods (e.g. majority vote or advanced decision-making algorithms) and prudently enforcing safety procedures that can enhance HRI. Furthermore, addition of hip muscle data or fusing data from other sources such as IMUs or force sensors are predicted to increase the accuracy. Deep learning method can be investigated to improve upon the C-ML in terms of generalization and adaptability, and enhance the prediction accuracy by providing transfer learning framework.

7. REFERENCES

[1] Andreas Meyer-Heim and Hubertus JA van Hedel, "Robot-assisted and computer-enhanced therapies for children with cerebral palsy: current state and clinical implementation," in *Seminars in pediatric neurology*. Elsevier, 2013, vol. 20, pp. 139–145.

[2] Luzheng Bi, Cuntai Guan, et al., "A review on emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration," *Biomedical Signal Processing and Control*, vol. 51, pp. 113–127, 2019.

[3] Bernard Hudgins, Philip Parker, and Robert N Scott, "A new strategy for multifunction myoelectric control," *IEEE transactions on biomedical engineering*, vol. 40, no. 1, pp. 82–94, 1993.

[4] Kevin Englehart, Bernard Hudgins, Philip A Parker, and Maryhelen Stevenson, "Classification of the myoelectric signal using time-frequency based representations," *Medical engineering & physics*, vol. 21, no. 6-7, pp. 431–438, 1999.

[5] Kevin Englehart, B Hudgin, and Philip A Parker, "A wavelet-based continuous classification scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, pp. 302–311, 2001.

[6] Kevin Englehart and Bernard Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE transactions on biomedical engineering*, vol. 50, no. 7, pp. 848–854, 2003.

[7] Ilja Kuzborskij, Arjan Gijsberts, and Barbara Caputo, "On the challenge of classifying 52 hand movements from surface electromyography," in *annual international conference of the IEEE engineering in medicine and biology society*, 2012, pp. 4931–4937.

[8] Manfredo Atzori, Arjan Gijsberts, Ilja Kuzborskij, Simone Elsig, Anne-Gabrielle Mittaz Hager, Olivier Deriaz, Claudio Castellini, Henning Müller, and Barbara Caputo, "Characterization of a benchmark database for myoelectric movement classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 1, pp. 73–83, 2014.

[9] Manfredo Atzori, Arjan Gijsberts, Claudio Castellini, Barbara Caputo, Anne-Gabrielle Mittaz Hager, Simone Elsig, Giorgio Giatsidis, Franco Bassetto, and Henning Müller, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Scientific data*, vol. 1, no. 1, pp. 1–13, 2014.

[10] Arjan Gijsberts, Manfredo Atzori, Claudio Castellini, Henning Müller, and Barbara Caputo, "Movement error rate for evaluation of machine learning methods for semg-based hand movement classification," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 22, no. 4, pp. 735–744, 2014.

[11] Ericka Janet Rechy-Ramirez and Huosheng Hu, "Bio-signal based control in assistive robots: a survey," *Digital Communications and networks*, vol. 1, no. 2, pp. 85–101, 2015.

[12] Abolfazl Mohebbi, "Human-robot interaction in rehabilitation and assistance: a review," *Current Robotics Reports*, pp. 1–14, 2020.

[13] Fan Zhang and He Huang, "Decoding movement intent of patient with multiple sclerosis for the powered lower extremity exoskeleton," in *annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, pp. 4957–4960.

[14] Nurhazimah Nazmi, Mohd Azizi Abdul Rahman, Shin-Ichiroh Yamamoto, Siti Anom Ahmad, Hairi Zamzuri, and Saiful Amri Mazlan, "A review of classification techniques of emg signals during isotonic and isometric contractions," *Sensors*, vol. 16, no. 8, pp. 1304, 2016.

[15] Kenneth R Lyons and Sanjay S Joshi, "A case study on classification of foot gestures via surface electromyography," in *Annu. Conf. Rehabil. Eng. Assist. Technol. Soc. Amer.*, 2015, pp. 1–5.

[16] Huong Thi Thu Vu, Dianbiao Dong, Hoang-Long Cao, Tom Verstraten, Dirk Lefeber, Bram Vanderborght, and Joost Geeroms, "A review of gait phase detection algorithms for lower limb prostheses," *Sensors*, vol. 20, no. 14, pp. 3972, 2020.

[17] Mads Jochumsen and Imran Khan Niazi, "Detection and classification of single-trial movement-related cortical potentials associated with functional lower limb movements," *Journal of Neural Engineering*, vol. 17, no. 3, pp. 035009, 2020.

[18] Blair Hu, Elliott Rouse, and Levi Hargrove, "Fusion of bilateral lower-limb neuromechanical signals improves prediction of locomotor activities," *Frontiers in Robotics and AI*, vol. 5, pp. 78, 2018.

[19] "Surface electromyography for the non-invasive assessment of muscles," http://www.seniam.org/.

[20] Angkoon Phinyomark, Pornchai Phukpattaranont, and Chusak Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

AN OPEN-SOURCE PLATFORM FOR COOPERATIVE, SEMI-AUTONOMOUS ROBOTIC SURGERY

Laura Connolly[1], *Anton Deguet* [2], *Kyle Sunderland*[1], *Andras Lasso*[1], *Tamas Ungi* [1], *John F. Rudan*[1], *Russell H. Taylor*[2], *Parvin Mousavi* [1], *Gabor Fichtinger* [1]

[1]Queen's University, Kingston, Ontario
[2]Johns Hopkins University, Baltimore, Maryland

ABSTRACT

Introduction: In this paper, we present and assess a proof of concept platform for semi-autonomous, cooperative robotic surgery. The platform is easily reproducible thanks to simple hardware components and open-source software. Moreover, the design accommodates open, soft tissue surgeries that recent advancements in surgical robotics do not generally focus on. **Methods:** The system is made up of an inexpensive robotic manipulator, a navigation system and a software interface. Accuracy measurement is performed on a rigid phantom that mimics the conditions of breast conserving surgery (BCS) as an example of a surgical use case. **Results:** The average target registration error (TRE) and fiducial registration error (FRE) of the system is within 1 mm. This indicates that the navigation system is sufficient for certain surgical applications such as BCS. The platform can also be easily replicated and used in a lab or home environment.

***Index Terms*—** Open source medical robotics, Semi-autonomous robotic surgery, Soft tissue surgery, Cooperative robotics.

1. INTRODUCTION

There has been a significant uptake in the adoption of surgical robotic systems in the last 20 years. In general, these systems are used to compensate for inaccuracies in human positioning and control of surgical instruments. Additionally, these systems can be used to improve dexterity in confined anatomical locations which can lead to less invasive surgery, smaller incisions and shorter recovery periods for patients [1]. Soft tissue surgery specifically, can greatly benefit from the addition of robotic guidance because the tissue is highly deformable which makes following a predefined surgical plan by hand, very challenging. However, as the field moves towards applying robotics to various surgical procedures, open, soft tissue surgery is still often overlooked.

Breast conserving surgery (BCS) is an example of such procedure. BCS involves a small excision to remove a tumor in the breast with as little skin and healthy tissue loss as possible. Many breast cancer patients prefer BCS to other treatment options like mastectomy (where the entire breast is removed), because it can prevent severe breast deformity and consequently, reconstructive surgery. As it is now, this procedure has approximately a 30 percent failure rate because precise tumor delineation by hand is extremely difficult as cancer lesions are often poorly defined, irregularly shaped and non-palpable [2]. Applying robotics to BCS could potentially increase the success rate of these procedures, by providing physical guidance as to where the margins of the tumor are. However, current surgical robots that are commercially available are too expensive and intrusive for wide spread surgical adoption in this use case [3]. Additionally, the actuators of existing devices are designed for small incision sites, unlike the open resection cavity in this type of surgery.

Existing teleoperated surgical robots typically make use of a custom manipulator (the follower) that is operated by a separate workstation (the leader) [4]. An example of this type of device is the Endo[PA]R system for minimally invasive heart surgery which is made up of a manipulator that is controlled with two separate leaders, namely two PHANToM haptic devices [5]. In this case, the manipulator is designed to accommodate the heavy resistant forces of the chest walls while the leader is just used for sensing the surgeons desired hand motion. In soft tissue surgery (as opposed to heart surgery), we realize that the tissue under exploration poses practically no resistant forces on the device, therefore, the strength of the leader itself should be sufficient for actuation. Under this assumption, we dissect this teleoperated systems approach by using the leader for surgical guidance, exclusively.

To achieve this, we make use of a cooperative control concept, first demonstrated by the Steady-Hand robot where the surgical instrument can be held by both a manipulator and a surgeon [6]. While the Steady-Hand robot uses a custom reversible chain-driven motor, applying this concept to soft-tissue surgery does not require the same level of precision control. With this in mind, our approach can be developed with an off-the-shelf robotic device that can achieve cooperative control with low cost and convenience.

Therefore, in this paper we present a semi-autonomous,

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

proof of concept platform for cooperative surgical guidance. To construct this system, we make use of inexpensive materials and open-source technology so the system is feasible for adoption in soft tissue and out-patient procedures. The accuracy of the system is evaluated using a phantom that is designed to mimic a real surgical work space. With this evaluation, we show that the platform can provide accurate tracking (within 1 mm) and discuss the necessary adaptations that should be made for application in BCS.

The paper is outlined as follows: Section 2 demonstrates the different components that are used to construct the platform, followed by the experiments that are performed to evaluate the system in Section 3. In Section 4 the results of these experiments are presented, alongside a discussion about potential clinical extension of the platform in BCS. Finally, Section 5 summarizes our findings and future work.

2. SYSTEM OVERVIEW

The platform we designed has three main components: the cooperative robotic manipulator, navigation system and software interface. The robotic manipulator is used for cooperative guidance and tool steering, while the navigation system provides supplemental information about the position of the soft tissue relative to the robot. The software interface provides visual feedback to the surgeon for navigation as well as a central interface for sending force commands to the manipulator. The following sections will describe each component in more detail.

2.1. Cooperative robotic manipulator

The robot we use in our system is the leader in a leader/follower configuration. The device itself is the Omni Bundle (Quanser, Markham, ON, Canada), the most recent iteration of the SensAble PHANToM Omni [7]. This is a 6 degree-of-freedom (DOF) robot with 3 actuated and 3 passive joints as well as built in haptics. The passivity in the wrist of the robot enables cooperative movement of the end-effector (EE). The robot also comes with a bracket that can be used to fix the wrist to the second-last joint to hold the EE at a specific orientation (Figure 2). This device is selected because of its low cost and small size when compared to similar haptic devices, both of which make it an ideal candidate for prototyping.

In this system the robot can provide both physical assistance by holding the surgical instrument as well as accurate tool tracking through the device's built in encoders. We can also introduce virtual constraints to perform tool steering in surgery, such as applying force whenever the surgeon approaches the boundaries of a tumor.

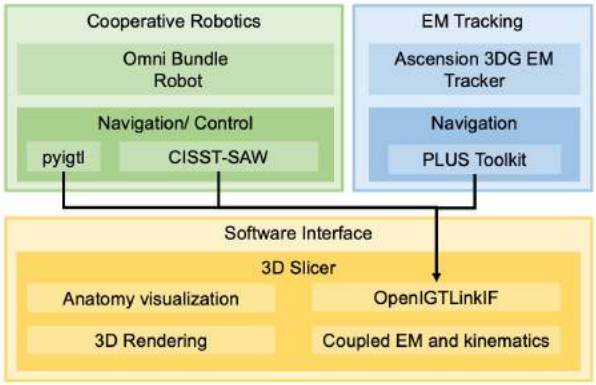


Fig. 1. Overview of platform components and software libraries used for integration.

2.2. Navigation

Alongside the manipulator tracking, the platform also incorporates electromagnetic (EM) navigation [2]. Supplemental EM tracking is necessary because we can use an EM sensor to provide a 6 DOF reference to the position of the soft tissue. For soft tissue guidance, this reference sensor is important because static preoperative images are not sufficient when the tissue is mobile during surgery. Moreover, as the system is cooperative, EM tracking gives the surgeon freedom to remove the device from the manipulator without losing navigational information. EM tracking was also chosen over optical tracking because it does not require a direct line of sight to target anatomy [8].

It has been previously demonstrated that EM tracking is feasible in a surgical environment, despite potential distortion factors [2], [9]. However, there has been minimal investigation into combined EM tracking and robotic guidance. For the purpose of this platform, we hypothesize that the addition of a small manipulator will not increase the tracking error of the EM system.

2.3. Software interface

Finally, the software interface for the platform was developed to fulfill our two functional requirements; 3D visualization of the manipulator and navigation data, and semi-autonomous control. The interface is created as an extension within the open-source, medical imaging platform, 3D Slicer (www.slicer.org) because it offers visualization features and developer tools through various external libraries. To use these visualization tools, the encoder values for each joint of the robot are streamed through the CISST-SAW library developed at Johns Hopkins (https://github.com/jhu-cisst/cisst-saw), via OpenIGTLink (www.openIGTLink.org) into 3D Slicer. Within the CISST-SAW library we make use of the sawSensablePhantom package that offers a direct interface

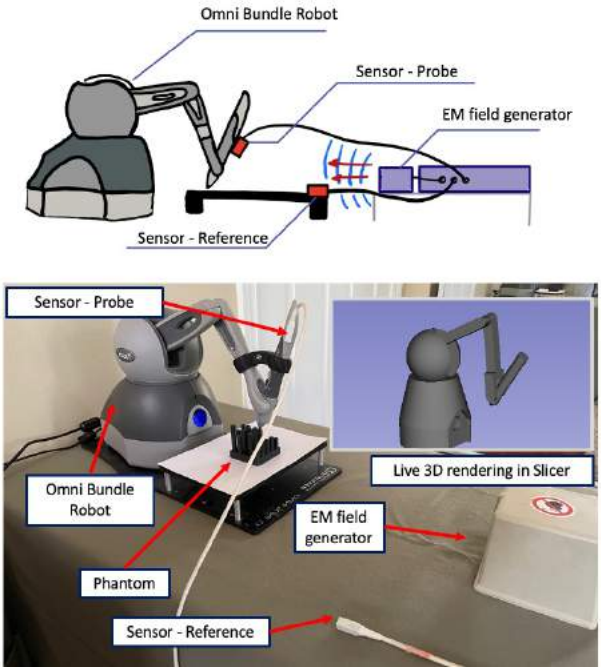


Fig. 2. Top - Diagram of experimental setup used for accuracy testing. Bottom - Experimental setup used for accuracy testing.

for the Omni Bundle robot, and modify the infrastructure of SlicerOpenIGTLink to receive the appropriate data streams. This presents an additional benefit because sawSensablePhantom can be setup in a Windows environment, whereas most robot infrastructure is executed using ROS which only runs on Unix-based platforms. The encoder values of each joint are then used to control an STL model of the robot in 3D, allowing for near real-time visualization of the manipulator relative to preoperative medical images.

Control of the device is facilitated in a similar way, using a library called pyigtl (https://github.com/lassoan/pyigtl). This library is used for sending messages to control the Omni Bundle via OpenIGTLink, again using sawSensablePhantom. For users, the software interface is equipped with buttons that can be used to move the EE incrementally in each direction or hold the manipulator in place to enforce the position of the EE. An overview of the entire system and the libraries used for each part of the integration is demonstrated in Figure 1.

3. EXPERIMENTS

During testing, one EM sensor was fixed to the table and one was fixed to the EE of the robot using hot glue. The EM field generator was then placed directly in front of the robot. This setup is demonstrated in Figure 2. We performed a pivot calibration to determine the EE tip to EM sensor coordinate system before each trial by detaching the bracket that secures

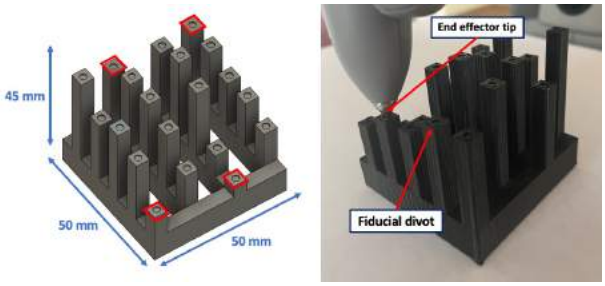


Fig. 3. Left - Phantom model in CAD with measurements. Registration points outlined in red. Right- 3D printed phantom.

the last two joints. To calibrate the manipulator, a homing process was done at startup using the 'inkwell' of the device.

3.1. Printed Phantom

For evaluating our system, a rigid phantom was 3D printed based on a similar study conducted by Kim et al. [10]. The phantom is a 50 x 50 mm base frame with 20 fiducials at different heights varying from 15 to 45 mm. In total, the fiducials on the phantom represent various fixed positions in a work space with a volume of 50 x 50 x 30 mm^3, which mimics that of a real breast cavity during tumor resection. The fiducials were placed at various heights to simulate a real tumor resection where the tool tip is moved in every direction within the cavity. A divot that provided a tight fit for the EE on the fiducial was also printed on each point to ensure that the locations that the system was being navigated to, represented the center of each fiducial (Figure 3).

The robot was guided to each fiducial and a control signal was used to hold the robot in place autonomously while the position in both coordinate systems was recorded. The 4 fiducials on the corners (outlined in red in Figure 3) were used for registration and the remaining 16 were used for testing. This same test was performed 9 times, re-calibrating each time, to assess the accuracy of both the EM and manipulator tracking.

A rigid, landmark registration was performed on the four registration points, to transform the EM and robotic coordinate systems to the ground truth positions of the fiducials based on the printed specifications of the phantom. The average fiducial registration error (FRE) of the registration points and target registration error (TRE) of the test points was then calculated. The FRE (1) represents the difference between the 4 transformed registration points in each coordinate system against the ground truth, while TRE (2) represents the difference between 16 transformed test points against the ground truth. These metrics are defined by,

$$FRE_i = |R\boldsymbol{f_1} + \boldsymbol{t} - \boldsymbol{f_2}| \tag{1}$$

$$TRE(\boldsymbol{p_2}) = |R\boldsymbol{p_1} + \boldsymbol{t} - \boldsymbol{p_2}| \tag{2}$$

such that $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$ are corresponding registration points, $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ are corresponding test points, R is the rotation matrix between each point pair, and $\boldsymbol{t}$ is the translation vector between each point pair [11] [12].

4. RESULTS AND DISCUSSION

Table 1 demonstrates the average TRE and FRE of the experiments, along with the recorded standard deviation, in each direction, across all 9 trials.

Table 1. Average and directional TRE, FRE and standard deviation for both EM and Robotic tracking.

EM Tracking	FRE (mm)	TRE (mm)
x	1.06 ± 0.41	0.86 ± 0.40
y	1.02 ± 0.46	1.21 ± 0.57
z	0.52 ± 0.20	0.51 ± 0.14
Average	$\mathbf{0.87 \pm 0.30}$	$\mathbf{0.86 \pm 0.27}$
Manipulator Tracking	FRE (mm)	TRE (mm)
x	1.39 ± 0.17	0.68 ± 0.14
y	0.64 ± 0.13	0.73 ± 0.25
z	0.54 ± 0.10	0.75 ± 0.13
Average	$\mathbf{0.86 \pm 0.11}$	$\mathbf{0.72 \pm 0.10}$

For both tracking modalities, the average TRE and FRE of the fiducials of the system is within 1 mm. For certain surgical procedures like breast cancer resection, where the acceptable margin of error is approximately 2 mm, this level of accuracy is sufficient. With this level of tracking, and the ability to send force commands to guide the manipulator in 3D Slicer, the platform fulfills the necessary control action and positioning requirements for the development of semi-autonomous control schemes. Additionally, the low TRE and FRE in both the EM and manipulator coordinate systems show that the robot does not pose any significant distortion on the EM field that is used for tracking. Although the effect of distortion posed by the surgical environment on EM tracking has been previously investigated, this is an important observation for continued investigation of combined EM navigation and robotics.

Extension to navigated BCS: The platform can be extended for navigation in BCS. BCS is a surgical procedure that already incorporates EM navigation [2]. The EM tracking helps the surgeon follow a predefined plan based on a tumor delineation that is preformed before the resection begins, with ultrasound guidance [2]. The tumor is used as a soft tissue reference point by placing a needle through the center and EM tracking the needle. With this, the system can provide the surgeon with visual feedback of the location of their tool relative to the margins of the tumor as they operate. Adding robotic guidance to this system would help the surgeon follow the plan more accurately by providing tactile feedback through the manipulator in addition to this visual feedback. Hands-on control can also be used to guide the surgeon along the borders of the tumor for careful resection.

As previously mentioned, the platform that we have presented is an appropriate test bed for in-lab development and research of semi-autonomous surgical robotic systems. To extend the system for clinical application, like BCS, would require a few modifications. One of the major changes would be swapping the low-cost manipulator with a larger robot that passes regulatory approval for intraoperative use. The next step would be attaching the resection device, to the EE of the robot and factoring the device offset into the EM pivot calibration and internal tracking of the manipulator. Ideally this would be done with a 3D printed clip that will allow for quick release of the surgical instrument if necessary. This is an additional benefit to this platform as many other surgical robots do allow for quick swapping or detachment of the instruments that are being controlled. The second modification would be performing sensor fusion with the existing EM data and the soft tissue reference provided by the tracked needle in the tumor.

5. CONCLUSION

A platform for semi-autonomous, cooperative robotics in soft tissue surgery is presented and evaluated in this study. The platform is designed using low-cost and open-source tools, which promote accessibility, flexibility and extension of robotic development to surgical areas that are currently under-investigated. To evaluate the system, a rigid phantom that represents the dynamic size and makeup of a surgical cavity is designed and used for testing. The system has an average FRE and TRE below 1 mm which indicates that it can provide acceptable tracking accuracy for surgical procedures, e.g. such as BCS.

This platform is an accurate, open source, low-cost, test bed that can be constructed in any environment. As the recent pandemic has caused many researchers to lose access to their lab space, this system can provide an alternative for at-home prototyping and testing. Beyond that, clinical deployment can be done by replacing the existing components with devices and technology that have already received regulatory approval. Future work involves adding open-source control schemes to the software interface for this system to execute hands-on and handheld surgical guidance.

ACKNOWLEDGMENT

Laura Connolly was supported by NSERC and CIHR. G. Fichtinger is supported as a Canada Research Chair. This work was funded, in part, by CANARIE's Research Software Program.

6. REFERENCES

[1] Brian S. Peters, Priscila R. Armijo, Crystal Krause, Songita A. Choudhury, and Dmitry Oleynikov, "Review of emerging surgical robotic technology," apr 2018.

[2] Gabrielle Gauvin, Caitlin T. Yeo, Tamas Ungi, Shaila Merchant, Andras Lasso, Doris Jabs, Thomas Vaughan, John F. Rudan, Ross Walker, Gabor Fichtinger, and Cecil Jay Engel, "Real-time electromagnetic navigation for breast-conserving surgery using NaviKnife technology: A matched case-control study," *The Breast Journal*, vol. 26, no. 3, pp. 399–405, mar 2020.

[3] Philipp Schleer, Sergey Drobinsky, Matias de la Fuente, and Klaus Radermacher, "Toward versatile cooperative surgical robotics: a review and future challenges," oct 2019.

[4] Max B. Schäfer, Kent W. Stewart, and Peter P. Pott, "Industrial robots for teleoperated surgery - A systematic review of existing approaches," *Current Directions in Biomedical Engineering*, vol. 5, no. 1, pp. 153–156, sep 2019.

[5] Hermann Mayer, István Nagy, Alois Knoll, Eva U. Schirmbeck, and Robert Bauernschmitt, "The Endo[PA]R system for minimally invasive robotic surgery," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2004, vol. 4, pp. 3637–3642.

[6] Russell Taylor, Pat Jensen, Louis Whitcomb, Aaron Barnes, Rajesh Kumar, Dan Stoianovici, Puneet Gupta, Zheng Xian Wang, Eugene DeJuan, and Louis Kavoussi, "Steady-hand robotic system for microsurgical augmentation," *International Journal of Robotics Research*, vol. 18, no. 12, pp. 1201–1210, jul 1999.

[7] Philipp Schleer, Philipp Kaiser, Sergey Drobinsky, and Klaus Radermacher, "Augmentation of haptic feedback for teleoperated robotic surgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, no. 3, pp. 515–529, mar 2020.

[8] Tamas Ungi, Andras Lasso, and Gabor Fichtinger, "Open-source platforms for navigated image-guided interventions," oct 2016.

[9] Elodie Lugez, Hossein Sadjadi, David R. Pichora, Randy E. Ellis, Selim G. Akl, and Gabor Fichtinger, "Electromagnetic tracking in surgical and interventional environments: usability study," *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, no. 3, pp. 253–262, mar 2015.

[10] Sungmin Kim, Youri Tan, Anton Deguet, and Peter Kazanzides, "Real-time image-guided telerobotic system integrating 3d slicer and the da vinci research kit," in *Proceedings - 2017 1st IEEE International Conference on Robotic Computing, IRC 2017*. may 2017, pp. 113–116, Institute of Electrical and Electronics Engineers Inc.

[11] J. Michael Fitzpatrick, "Fiducial registration error and target registration error are uncorrelated," in *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*. feb 2009, vol. 7261, p. 726102, SPIE.

[12] J. Michael Fitzpatrick, Jay B. West, and Calvin R. Maurer, "Predicting error in rigid-body point-based registration," *IEEE Transactions on Medical Imaging*, vol. 17, no. 5, pp. 694–702, 1998.

IMPROVING A USER'S HAPTIC PERCEPTUAL SENSITIVITY BY OPTIMIZING EFFECTIVE MANIPULABILITY OF A REDUNDANT USER INTERFACE

Teng Li[1], *Ali Torabi*[1], *Hongjun Xing*[2], *and Mahdi Tavakoli*[1]*

[1]University of Alberta, Edmonton T6G 1H9, Alberta, Canada.
[2]Harbin Institute of Technology, Harbin 150001, China.

ABSTRACT

Human perceptual sensitivity of various types of forces, *e.g.*, stiffness and friction, is important for surgeons during robotic surgeries such as needle insertion and palpation. However, force feedback from robot end-effector is usually a combination of desired and undesired force components which could have an effect on the perceptual sensitivity of the desired one. In presence of undesired forces, to improve perceptual sensitivity of desired force could benefit robotic surgical outcomes. In this paper, we investigate how users' perceptual sensitivity of friction and stiffness can be improved by taking advantage of kinematic redundancy of a user interface. Experimental results indicated that the perceptual sensitivity of both friction and stiffness can be significantly improved by maximizing the effective manipulability of the redundant user interface in its null space. The positive results provide a promising perspective to enhance surgeons' haptic perceptual ability by making use of the robot redundancy.

Index Terms— Haptic Perception, Kinematic Redundancy, Effective Manipulability, Viscous Friction, Stiffness

1. INTRODUCTION

Discriminating the properties of soft tissue, such as different levels of stiffness, is important for surgeons to perform some surgical procedures like needle insertion and palpation [1, 2]. In robotic surgery, force feedback delivered to the surgeons from the robot end-effector is usually a combination of several force components including such as soft tissue stiffness and friction, robot inertia, and joint friction. In this case, the desired force, *e.g.*, tissue stiffness, could easily be affected by other undesired ones [3, 4].

This research is supported in part by the Canada Foundation for Innovation (CFI) under grants LOF 28241 and JELF 35916, in part by the Government of Alberta under grants IAE RCP-12-021 and EDT RCP-17-019-SEG, in part by the Government of Alberta's grant to Centre for Autonomous Systems in Strengthening Future Communities (RCP-19-001-MIF), in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under grants RGPIN-2019-04662 and RGPAS-2019-00106, and in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under grant RTI-2018-00681. *Correspondence: mahdi.tavakoli@ualberta.ca

In the presence of undesired forces, perceptual sensitivity of the desired one can be largely affected. For example, in surgical procedure of needle insertion, tip force is often combined with needle shaft friction and could be masked by each other [2, 5], which makes the discrimination of either of them more difficult. As a consequence, the perceptual sensitivity of the desired force will decrease as the magnitude of the undesired one increases.

Improving the perceptual sensitivity of desired force in the presence of undesired ones could be beneficial to the surgeons' performance as well as the robotic surgical system. With high haptic perceptual sensitivity, surgeons can accurately localize a lesion and judge the healthy status of target tissue [6]. For some surgeries, the haptic perceptual sensitivity could be critical. For example, in retinal microsurgery, only about 20% of events can be detected in which the tiny forces are around $7.5mN$ [7]. Just noticeable difference (JND) and Weber fraction (WF) are two commonly used characteristics to measure human perceptual sensitivity [8, 9].

There are some methods have been developed to enhance users' perceptual sensitivity, such as scaling force feedback and developing new tools. Scaling force feedback is a commonly used method to better meet human perceptual ability, especially in teleoperation systems [1]. Considering that the desired force is usually mixed with noises, scaling force feedback will scale all noises simultaneously. Besides, scaling force may distort users' feeling and make it unreal.

De Lorenzo *et al.* [5] introduced a new device, a robotic coaxial needle insertion assistant, to enhance human perceptual sensitivity. The device is able to separate the needle tip force and needle shaft-tissue friction force during needle insertion. With this device, the undesired forces can be filtered out, thus enhancing the perceptual sensitivity of the desired one. However, a new device cannot be easily introduced into the operating room due to various regulatory approvals that it must go through.

Kinematic redundancy has been used to improve task performance on modeled soft tissue stiffness discrimination by comparing redundant and non-redundant robot [10]. The advantage of this method is that it is making use of the intrinsic property of redundant robots, *i.e.*, having a larger effective manipulability (EM) than non-redundant robots, and without

additional costs. Our previous work in [10] was focusing on general redundant robots. Here in this work we will narrow down to focus on one specific redundant robot and investigate how haptic perceptual sensitivity can be affected by different methods of optimization in the redundant robot's null space.

In this paper, we are considering scenarios of tangential palpation (friction discrimination) and needle insertion (stiffness discrimination) where both desired and undesired forces will be in presence. Please note that, we will not pay too much attention on the potential masking effect of the undesired force in this paper. Instead, we will focus on taking the intrinsic advantage of kinematic redundancy to investigate the following two questions,

1. How perceptual sensitivity of friction and stiffness will be affected by different methods of optimizing the effective manipulability (EM) of a redundant robot?

2. Is there any trade-off effect on the haptic perceptual sensitivity when optimizing the EM to be isotropic?

Our hypotheses are that, the perceptual sensitivity of both friction and stiffness can be improved by maximizing the EM along the movement direction, and there is a trade-off effect for isotropic condition.

The remaining part of this paper is organized as follows: Section 2 describes the methods in detail including apparatus, cost function and control law, participants and experimental conditions. Section 3 presents experimental results and discussions. Section 4 remarks on our conclusions.

2. METHODS

2.1. Apparatus

A custom 4-degree-of-freedom (DOF) redundant planar haptic device including two robots, as shown in Figure 1, was employed in our experiments. The first base robot was a 2-DOF planar Rehabilitation robot (Quanser Inc., Markham, ON, Canada) while the second one came from a PHANToM 1.5A (Geomagic Inc., Morrisville, NC, USA). The 4-DOF robot was controlled via interface of MATLAB/Simulink (R2017a, MathWorks Inc., Natick, MA, USA) with Quarc real-time control software (Quanser Inc., Markham, ON, Canada). The control rate of the experiment was 1000 Hz.

2.2. Cost function and control law

The effective manipulability (EM), denoted as ρ in Eqn.(1), is commonly used to describe robot manipulability along a specified movement direction. Here we took it as our cost function to optimize the EM along a specified direction u via the internal motion of the redundant robot in its null space.

$$\rho = (u^T (JJ^T)^{-1} u)^{-1/2} \tag{1}$$

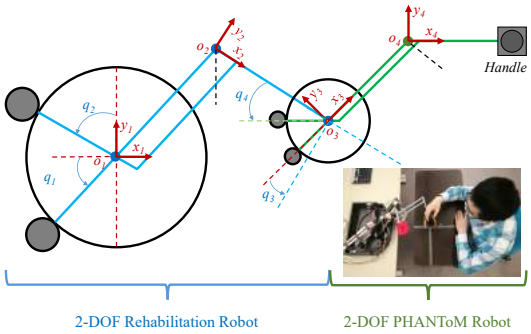


Fig. 1: Sketch of the 4-DOF robot and experimental scenario.

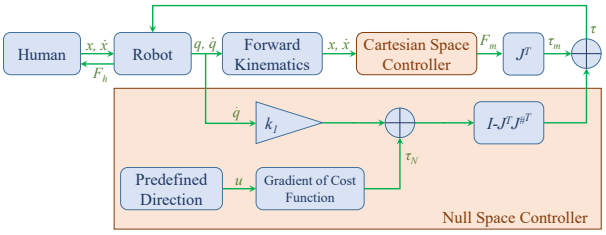


Fig. 2: Control diagram for the 4-DOF robot.

where J is the Jacobian matrix of the 4-DOF robot, and u is the specified movement direction. The velocity manipulability ellipsoid $M = JJ^T$ is included inside Eqn.(1).

The control diagram for the 4-DOF robot is shown in Figure 2, in which the null space controller [11] is defined by

$$\tau = J^T F_m + (I - J^T J^{\#^T})(\tau_N - k_1 \dot{q}) \tag{2}$$

where τ is the joint torque vector for generating the robot end-effector force F_m, and τ_N is related to the gradient of the cost function Eqn.(1) which will be projected into the robot null space by a projector $(I - J^T J^{\#^T})$. The parameter of k_1 is a suitable positive constant damping gain for stabilizing the system while $J^{\#}$ is the generalized inverse Jacobian.

The Cartesian space controller for the primary task was modeled as a spring-damper model, *i.e.*, a virtual wall with friction, as follows

$$F_m = k_D(\dot{x}_d - \dot{x}) + k_P(x_d - x) \tag{3}$$

where k_P is the spring coefficient, k_D is the damping coefficient, x and $\dot{x}$ are the real-time end-effector position and velocity respectively, while x_d and $\dot{x}_d$ are the desired end-effector position and velocity respectively. In this paper, we modeled the tissue friction as viscous damping and modeled the tissue stiffness as spring stiffness. By tuning the damping coefficient k_D and the spring coefficient k_P, the tissue friction and stiffness can be adjusted respectively.

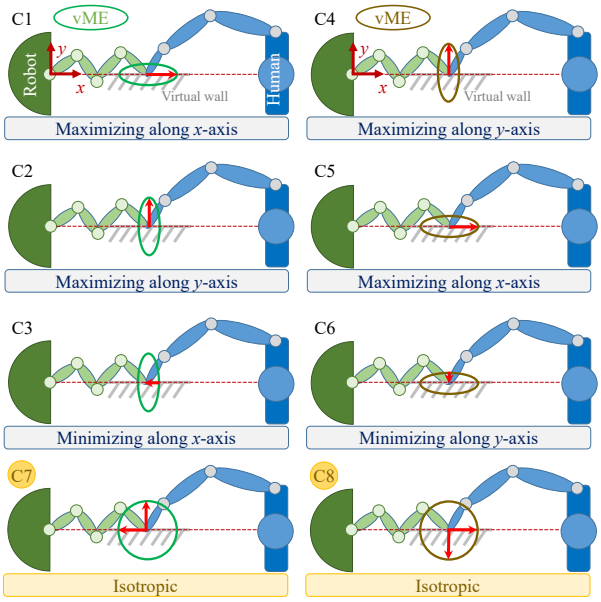


Fig. 3: Illustration of experimental conditions (top view). The red arrows represent the optimization directions. vME, means velocity manipulability ellipsoid.

2.3. Participants and experimental conditions

Six participants were employed for the experiments. The experiments were approved by the Health Research Ethics Board (HREB) at University of Alberta under study ID MS3_Pro00057919. Please note that, due to the COVID-19 pandemic, all experiments involving human subjects were suspended at University of Alberta, therefore all participants were played by the first author.

In total of three experiments including eight conditions (C1~C8) were designed as shown in Figure 3, and different conditions indicated different optimization methods. Experiment-1 (C1,C2,C3) was friction discrimination task where the directions of friction and stiffness were orthogonal to each other as illustrated in Figure 4. Experiment-2 (C4,C5,C6) was stiffness discrimination task where the directions of friction and stiffness were parallel to each other. Experiment-3 (C7,C8) included two isotropic conditions which can be viewed as the extended condition for Experiment-1 and Experiment-2 respectively.

Based on two alternative forced choice (2AFC) method [12, 13], in each trial of all experiments, participants were required to discriminate two stimuli (one reference and one comparison, sequentially and randomly presented), then answered a predefined question of *"whether the second tissue friction/stiffness is higher than the first one?"* by typing in number 1 ("yes") or number 0 ("no"). Nine friction/stiffness levels yielded 90 trials in total (9 *pair* $\times$ 10 *repetition*) for each condition each participant.

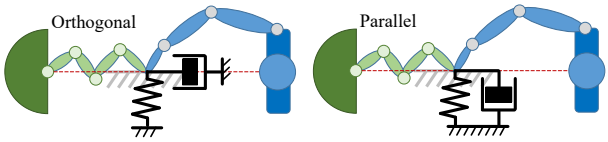
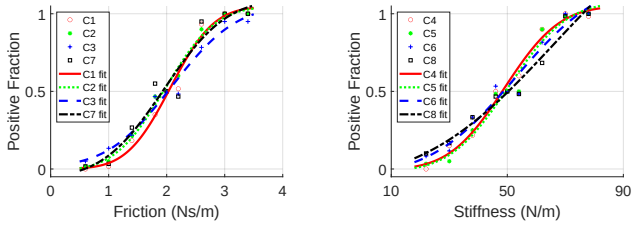


Fig. 4: Illustration of relative directions between stiffness and friction. "Orthogonal" is for friction discrimination task while "Parallel" is for stiffness discrimination task.



(a) Friction (b) Stiffness

Fig. 5: Positive fraction with respect to friction and stiffness.

3. RESULTS AND DISCUSSIONS

3.1. Performance metrics

For all experiments, we employed the just noticeable difference (JND) and the Weber Fraction (WF) as our major human performance metrics. The JND describes the minimum differences that have to be made between the comparison stimulus and the reference stimulus in order to perceive a noticeable change for the human. The WF describes the percentage difference in stimulus strength with respect to the reference stimulus that is just noticeable [10].

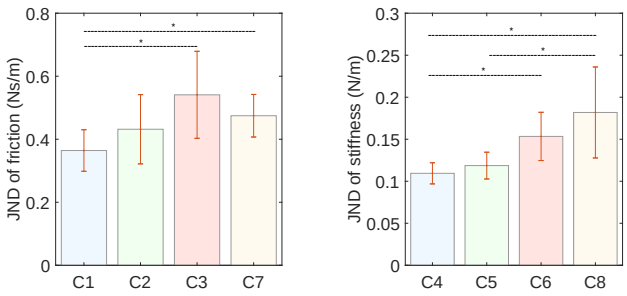
Using method of Weibull function [10, 14], the fitted psychometric functions based on all pooled data were obtained and shown in Figure 5. The JND and WF were calculated using commonly used method [8] and listed in Table 1.

3.2. Experiment-1 & Experiment-2

In order to investigate the perceptual sensitivity of friction and stiffness, we conducted Experiment-1 and Experiment-2 respectively. Experiment-1 of friction discrimination task can be taken as a mimic scenario of tangential palpation where the directions of friction and stiffness were *orthogonal* to each other. Experiment-2 of stiffness discrimination task can be taken as a mimic scenario of needle insertion where the directions of friction and stiffness were *parallel* to each other [5].

The results of Experiment-1 & Experiment-2 were shown in Table 1 and Figure 6. For simplicity, we also included the results of Experiment-3 (C7,C8) in the same table and figure.

By comparing the three conditions of C1,C2,C3 as well as the three conditions of C4,C5,C6 respectively in Table 1 and Figure 6, we can find that maximizing the EM along the



(a) Friction (b) Stiffness

Fig. 6: JND of friction and stiffness.

Table 1: Summary of JND and WF in each condition.

Friction Task			Stiffness Task		
Cond.	JND	WF	Cond.	JND	WF
C1	0.3642	0.1819	C4	0.1094	0.2282
C2	0.4317	0.2199	C5	0.1187	0.2399
C3	0.5410	0.2767	C6	0.1534	0.3184
C7	0.4746	0.2474	C8	0.1819	0.3731

movement direction (C1/C4) will result significantly higher perceptual sensitivity of friction/stiffness (*i.e.*, lower JND and WF) than minimizing it (C3/C6) (see Table 2 for p-values).

For C2 (or C5), it was expected to have similar results to C3 (or C6) while different results from C1 (or C4), but that was not the case. There were no significant difference between C2 and C1 or C3 in friction discrimination task, also no significant difference between C5 and C4 or C6 in stiffness discrimination task. The reason was that they were using different cost functions for the optimization after realizing the specific cost functions, which made them be not comparable.

The results here indicated that, by maximizing the EM along the movement direction, the participants' perceptual sensitivity of both friction and stiffness can be significantly improved in terms of JND and WF.

3.3. Experiment-3: Isotropic conditions

Experiment-3 (C7,C8) included two isotropic conditions where the EM was set to be isotropic rather than maximizing/minimizing it. The goal here was to investigate whether the isotropic conditions (C7,C8) will have a trade-off effect on perceptual sensitivity when comparing to condition of maximizing (C1,C4) and minimizing (C3,C6) EM.

By comparing C7 with C1,C3 in the friction discrimination task in Table 1 and Figure 6, we can find that there was only significant difference for C7 with C1 but not with C3. Also, the isotropic condition (C7) seems to have a trade-off performance compared to condition of maximizing (C1) and minimizing (C3) EM in terms of numerical JND and WF.

However, this was not true for the stiffness discrimination task. The isotropic condition (C8) had the lowest sensitivity of stiffness (*i.e.*, the highest JND and WF) compared to condition of maximizing (C4) and minimizing (C6) EM in terms of numerical JND and WF. For statistical analysis, there was only significant difference for C8 with C4 but not with C6.

There was no any trace of trade-off effect for C8 even numerically like observed in C7. This could be caused by masking effect in the stiffness discrimination task since the directions of friction and stiffness were parallel to each other. However, further experiments were required before drawing any conclusion about isotropic condition and trade-off effect.

Table 2: Summary of p-values of t-test for paired-sample.

	C2	C3	C7		C5	C6	C8
C1	0.2257	0.0178*	0.0168*	C4	0.3575	0.0104*	0.0332*
C2	-	0.1603	0.4343	C5	-	0.0529	0.0291*
C3	-	-	0.3152	C6	-	-	0.4121

Note: * for significance level under 5%.

3.4. Limitations

The WF of friction obtained in our friction discrimination task was in a normal range like that shown in literature (around 0.23) [15]. But there was relatively larger difference between the WF obtained in our stiffness discrimination task (around 0.16 in the literature). This difference could be probably caused by the potential masking effect which resulted in larger values of WF and JND of stiffness in our experiment.

The main limitation of this paper was the small participants pool and potential bias since all experiments were performed by the first author. In future work, we will employ more participants to increase individual diversity and eliminate potential bias.

4. CONCLUSION

Haptic perceptual sensitivity is a beneficial factor for surgeons accurately conducting many surgical tasks like suturing and palpation. In this paper, we experimentally showed that the haptic perceptual sensitivity of friction and stiffness can be improved in terms of just noticeable difference (JND) and Weber Fraction (WF) by appropriately optimizing the effective manipulability (EM) of a redundant robot.

This paper provided a preliminary but promising result to improve haptic perceptual sensitivity by taking advantage of kinematic redundancy. In future work, we will investigate how masking effect will influence the haptic perceptual sensitivity, as well as whether the same optimization approach can also benefit the haptic perceptual sensitivity of other types of forces such as torque and inertia.

5. REFERENCES

[1] Leonardo Meli, Claudio Pacchierotti, and Domenico Prattichizzo, "Experimental evaluation of magnified haptic feedback for robot-assisted needle insertion and palpation," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 13, no. 4, pp. e1809, 2017.

[2] Gourishetti Ravali and Muniyandi Manivannan, "Haptic feedback in needle insertion modeling and simulation," *IEEE reviews in biomedical engineering*, vol. 10, pp. 63–77, 2017.

[3] George A Gescheider, SJ Bolanowski Jr, and Ronald T Verrillo, "Vibrotactile masking: Effects of stimulus onset asynchrony and stimulus frequency," *The Journal of the Acoustical Society of America*, vol. 85, no. 5, pp. 2059–2064, 1989.

[4] Markus Rank, Thomas Schauß, Angelika Peer, Sandra Hirche, and Roberta L Klatzky, "Masking effects for damping jnd," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 2012, pp. 145–150.

[5] Danilo De Lorenzo, Yoshihiko Koseki, Elena De Momi, Kiyoyuki Chinzei, and Allison M Okamura, "Coaxial needle insertion assistant with enhanced force feedback," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 379–389, 2012.

[6] Dangxiao Wang, Siming Zhao, Teng Li, Yuru Zhang, and Xiaoyan Wang, "Preliminary evaluation of a virtual reality dental simulation system on drilling operation," *Bio-medical materials and engineering*, vol. 26, no. s1, pp. S747–S756, 2015.

[7] Puneet K Gupta, Pahick S Jensen, and Eugene de Juan, "Surgical forces and tactile perception during retinal microsurgery," in *International conference on medical image computing and computer-assisted intervention*. Springer, 1999, pp. 1218–1225.

[8] Netta Gurari, Katherine J Kuchenbecker, and Allison M Okamura, "Stiffness discrimination with visual and proprioceptive cues," in *World Haptics 2009-Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, 2009, pp. 121–126.

[9] Bing Wu, Roberta L Klatzky, and Ralph L Hollis, "Force, torque, and stiffness: Interactions in perceptual discrimination," *IEEE transactions on haptics*, vol. 4, no. 3, pp. 221–228, 2011.

[10] Ali Torabi, Mohsen Khadem, Kourosh Zareinia, Garnette Roy Sutherland, and Mahdi Tavakoli, "Application of a redundant haptic interface in enhancing soft-tissue stiffness discrimination," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1037–1044, 2019.

[11] Oussama Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, 1987.

[12] Marcia O'Malley and Michael Goldfarb, "The effect of force saturation on the haptic perception of detail," *IEEE/ASME transactions on mechatronics*, vol. 7, no. 3, pp. 280–288, 2002.

[13] Martin Grunwald, *Human haptic perception: Basics and applications*, Springer Science & Business Media, 2008.

[14] Felix A Wichmann and N Jeremy Hill, "The psychometric function: I. fitting, sampling, and goodness of fit," *Perception & psychophysics*, vol. 63, no. 8, pp. 1293–1313, 2001.

[15] Alejandro F Azocar, Amanda L Shorter, and Elliott J Rouse, "Damping perception during active ankle and knee movement," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 2, pp. 198–206, 2019.

TOWARD SEMI-AUTONOMOUS STIFFNESS ADAPTATION OF PNEUMATIC SOFT ROBOTS: MODELING AND VALIDATION

Majid Roshanfar[1], *Javad Dargahi*[1], *Amir Hooshiar*[2]

[1] Department of Mechanical Engineering, Concordia University, Montreal, QC Canada
[2] Department of Surgery, McGill University, Montreal, QC Canada

ABSTRACT

The constant stiffness of the soft medical robots imposes a cap on their force transmission capacity. To address this limitation, the objective of this study was to investigate the effects of chamber pressure on the stiffness of pneumatic soft robots. To this end, a single-chamber pneumatic soft robot was designed and fabricated. Afterward, a mechanistic model of the robot under external force and chamber pressure was developed. The model was solved as an initial value problem with homogeneous Neumann and Dirichlet boundary conditions. Comparison of the theoretical findings with experimental results for tip displacement and stiffness showed similar trends with a maximum error of 8.7%. The findings confirmed the feasibility of stiffness adaptation through chamber pressure regulation.

Index Terms— Soft robot, pneumatic actuation, Cosserat rod model, stiffness adaptation, surgical robot.

1. INTRODUCTION

1.1. Background

Soft robots have favorable properties for applications in minimally invasive surgery (MIS). However, there is a dilemma in the usability of soft surgical robots. Soft surgical robots are usually introduced percutaneously or through body orifices and are steered toward the target anatomy. Therefore, their low stiffness is desirable for steerability. However, at the target they are intended to perform a specific task, e.g., ablation [1], which requires force transmission to the environment. On the other hand, the force transmission capacity is directly related to the stiffness of the soft robot. Therefore, there is a compromise between the deformability and force transmission capacity of the soft robots in MIS procedures.
In this regard, the majority of the currently proposed soft MIS robots are of constant stiffness inherited from their material properties and their structural design [1]. The main limitation of the constant stiffness soft robots is that they possess a pre-determined maneuverability and force transmission range. For example, a highly flexible robot can be safely maneuvered through the lumens but may not be of enough force capacity, while a highly stiff robot can apply relatively high force but may not conform easily in highly tortuous trajectories. Therefore, having a soft robot with controllable stiffness (for intraoperative adaptability) is of high clinical importance. In fact, development of such a soft robot facilitates developing control frameworks toward semi-autonomous stiffness modulation for better intraoperative adaptation to task-specific requirements, e.g., force capacity. The motivation of this study was to propose and validate a mechanistic model to investigate the effects of chamber pressure on the stiffness of the pneumatic-driven soft robots. This model will facilitate future research toward semi-autonomous stiffness adaptation framework for interventional procedures using soft surgical robots.

1.2. Related Studies

Various soft robotic MIS systems have been proposed for cardiovascular interventions [2, 3], endoscopy [4], drug delivery [5], and general surgery applications [6]. Theoretically, soft robots possess infinite degrees of freedom (DoF). However, to simplify the problem, the soft robot's deformation has been modeled as a curve [7, 8, 9], or a set of small rigid segments with flexible joints [10]. Also, piecewise constant curvature (PCC) model has been widely used in the literature [11]. More recently, the Cosserat rod model has been adopted for modeling soft robots with MIS applications [12]. This method treats the small and large deformations of soft robots with a unified formulation, thus, simplifies the model derivation. Nevertheless, to the best of the author's knowledge, the effects of the internal pressure on the variation of the structural stiffness of the soft pneumatic robots have not been investigated yet.

1.3. Contributions

The main contributions of this study were: 1) modeling the deformation of a single-chamber pneumatic driven soft flexure using Cosserat rod model, 2) solution of the Cosserat model for a given tip force as an initial value problem (IVP),

This research was supported by the Natural Science and Engineering Research Council (NSERC) of Canada through NSERC CREATE Grant for Innovation-at-the-Cutting-Edge (ICE), Concordia University, and McGill University, Montreal, QC, Canada.

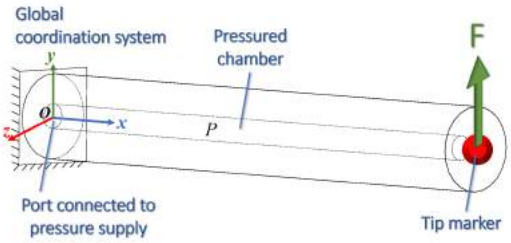


Fig. 1: Schematic initial shape of the soft robot with internal pressure P an external tip force $\mathbf{F}$.

3) validation of the proposed model through experimentation, 4) demonstrating the feasibility of stiffness modulation by changing the chamber pressure.

2. MECHANISTIC MODELING

In this section, first, the kinematics and force balance equations are derived. Afterward, the constitutive equation was derived to relate the forces and moments to kinematic displacements. Finally, the initial value problem (IVP) solution schema is presented, and the stiffness derivation from the tip force is provided.

2.1. Kinematics

Fig.1 depicts a hollow soft robot with a cross-sectional area A in its initial shape, subjected to an internal pressure P and an external tip force $\mathbf{F}$. The soft robot was also subjected to distributed gravitational force (weight). Each point on the backbone is parameterized by an arc parameter $s \in [0, L]$ and a locally orthonormal frame $\mathbf{R}(s)$ [11]. The position of any point on the arc, with respect to the base of the arc at a distance s, was defined by the position vector $\mathbf{p}(s)$. With the presented shape parameterization, the extension and shear strains along the backbone, $\mathbf{v}(s)$, are [13]:

$$\mathbf{v}(s) = \mathbf{R}^T(s)\frac{\partial \mathbf{p}(s)}{\partial s}, \tag{1}$$

while the bending and torsion strains, **v(s)**, are [13]:

$$\mathbf{u}(s) = \left(\mathbf{R}^T(s)\frac{\partial \mathbf{R}(s)}{\partial s}\right)^{\vee}, \tag{2}$$

where $(.)^{\vee}$ is the vee-operator, a mapping for $\mathfrak{so}(3)$ to $\mathbb{R}^3$ [14]. The kinematic equations relate the internal strains throughout the length of the robot to its parameterized shape.

2.2. Conservation of Momentum

Fig.2 depicts the free-body diagram of an arbitrary infinitesimally small element along the length of the soft robot. From a

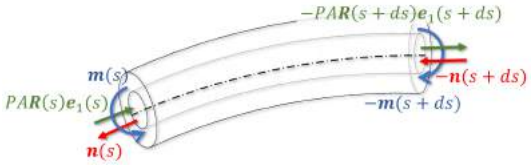


Fig. 2: Free-body diagram of an infinitesimally small element along the length of the soft robot.

mechanical point of view, the chamber pressure causes a constant longitudinal tensile force along the soft robot's length with a magnitude of PA, where A is the cross-sectional area of the robot perpendicular to its backbone. It is noteworthy that the internal pressure has symmetry about the longitudinal axis of the soft robot therefore it does not affect the force distribution in the perpendicular planes to the longitudinal axis. Using the fundamental Cosserat rod theory the quasi-static balance equations of the soft robot were obtained as [13]:

$$\frac{\partial \mathbf{p}(s)}{\partial s} = \mathbf{R}(s)\mathbf{v}(s), \tag{3}$$

$$\frac{\partial \mathbf{R}(s)}{\partial s} = \mathbf{R}(s)(\mathbf{u}(s))^{\wedge}, \tag{4}$$

$$\frac{\partial \mathbf{n}(s)}{\partial s} = -\rho A\mathbf{g} - PA\mathbf{e}_3(s), \tag{5}$$

$$\frac{\partial \mathbf{m}(s)}{\partial s} = -\left(\frac{\partial \mathbf{p}(s)}{\partial s}\right)^{\wedge}\mathbf{n}(s), \tag{6}$$

where $\mathbf{n}(s)$ and $\mathbf{m}(s)$ are the internal force and moment vectors in the global coordination system, P is the chamber pressure, ρ is the material density (constant), A is the cross-sectional area of the soft robot, $\mathbf{g} = \begin{pmatrix}0 & 0 & -9.81\end{pmatrix}^T$ is the gravity vector, $\mathbf{e}_1$ is the first unit vector of $\mathbf{R}(s)$ (tangential to the backbone), and $(.)^{\wedge}$ is the hat-operator, a mapping from $\mathbb{R}^3$ to $\mathfrak{so}(3)$ such that $((.)^{\wedge})^{\vee} = (.)$[14]. In fact, Eq. 1–6 describe the nonlinear state-space representation of the the soft robot's mechanics with six state variables, i.e., $\begin{pmatrix}\mathbf{u}(s) & \mathbf{v}(s) & \mathbf{p}(s) & \mathbf{R}(s) & \mathbf{m}(s) & \mathbf{n}(s)\end{pmatrix}$.

2.3. Constitutive Equations

Typically, the soft robots are made of hyperelastic elastomers. The mechanical properties of hyperelastic materials changes with local stretches, however for a given stretch (at any s) the tangent moduli $\mathbf{K}_{se}$ and $\mathbf{K}_{bt}$ represent the mechanical stiffness for unit length in shear and elongation momentarily. To calculate the tangent moduli, first a two-term Mooney-Rivlin (2MR) constitutive model for the material behavior of the soft robot was assumed. Eq. 7 represents the 2MR model for uniaxial elongation:

$$T_{11} = 2(c_{01} + \frac{c_{10}}{\lambda})(\lambda^2 - \lambda^{-1}), \tag{7}$$

with T_{11} longitudinal nominal stress, λ longitudinal stretch, c_{01} and c_{10} as material constants. Moreover, the initial shear modulus $G_\circ$ and initial Hooke's modulus $E_\circ$ of 2MR material is obtained by:

$$G_\circ = 2(c_{01} + c_{10}), \tag{8}$$

$$E_\circ = 2G_\circ(1 + \vartheta), \tag{9}$$

with ϑ as the Poisson's ratio that is ≈ 0.5 for near-incompressible elastomers. Based on the Cosserat rod model, the basic linear elastic constitutive equations are [13]:

$$\mathbf{n}(s) = \mathbf{R}(s)\mathbf{K}_{se}\Big(\mathbf{v}(s) - \mathbf{v}^\star(s)\Big), \tag{10}$$

$$\mathbf{m}(s) = \mathbf{R}(s)\mathbf{K}_{bt}\Big(\mathbf{u}(s) - \mathbf{u}^\star(s)\Big), \tag{11}$$

where, $(.)^\star$ refers to the state variables before deformation (initial state). Assuming, an initially straight soft robot extended along the global $x-$axis, $\mathbf{v}^\star(s) = \begin{pmatrix}1 & 0 & 0\end{pmatrix}^T$ and $\mathbf{u}^\star(s) = \mathbf{0}$. Also, substituting the derived shear and Hooke's moduli the tangent stiffness matrices were obtained as:

$$\mathbf{K}_{se} = \text{diag}\begin{pmatrix}E_\circ A & G_\circ A & G_\circ A\end{pmatrix}, \tag{12}$$

$$\mathbf{K}_{bt} = \text{diag}\begin{pmatrix}2G_\circ I & E_\circ I & E_\circ I\end{pmatrix}, \tag{13}$$

where I is the second moment of inertia of the soft robot's cross-section perpendicular to the backbone. The soft robot was also subjected to homogenous Dirichlet and Neumann boundary conditions at $s = 0$ that were formulated as:

$$\mathbf{p}(s)|_{s=0} = \begin{pmatrix}0 & 0 & 0\end{pmatrix}^T, \tag{14}$$

$$\mathbf{u}(s)|_{s=0} = \begin{pmatrix}1 & 0 & 0\end{pmatrix}^T, \tag{15}$$

$$\mathbf{R}(s)|_{s=0} = \mathbf{I}_{3\times 3}, \tag{16}$$

$$\mathbf{v}(s)|_{s=0} = \mathbf{0}. \tag{17}$$

Also, we assumed that the reaction forces at $s = 0$ were available by reading the force and moments from a six-DoF force/torque sensor placed at $s = 0$. Therefore, two addition boundary conditions were formulated as:

$$\mathbf{n}(s)|_{s=0} = -\mathbf{n}_\circ, \quad \mathbf{m}(s)|_{s=0} = -\mathbf{m}_\circ, \tag{18}$$

where $\mathbf{n}_\circ$ and $\mathbf{m}_\circ$ were the force and torque vectors in global coordination system, directly measured at $s = 0$.

2.4. Solution Schema

In order to find the deformation of the soft robot, initially the constitutive equations were substituted into the force and moment balance equations. Afterwards, by assuming *a-priori* knowledge of the reaction forces and moments at the $s = 0$, the system of nonlinear differential equations (Eq. 1–6) were integrated using 4-th order Runge-Kutta (RK4) method with a step-size of $\delta s = \frac{L}{100}$. Table 1 summarizes the model parameters used in the solution. Also, Fig. 3 depicts the deformation of the soft robot under its weight and a tip force of 30 mN in +z-direction and various chamber pressures. As depicted, increasing the internal pressure decreased the tip displacement. In other words, it showed that increasing the internal pressure had increased the bending stiffness of the soft robot.

Table 1: Model parameters of the prototyped soft robot.

Parameter	Length L (mm)	Outer Dia. D_o (mm)	Inner Dia. D_i (mm)	2MR Constants c_{01} (kPa)	2MR Constants c_{10} (kPa)	Density ρ ($\frac{g}{cc}$)
	85	12	3.5	277	-209	1.04

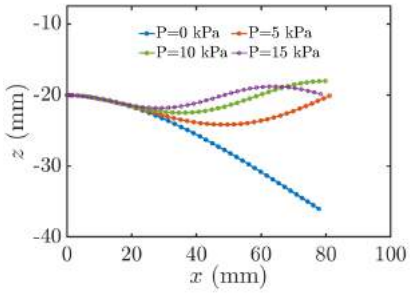


Fig. 3: Deformation of the soft robot under its weight and a tip force of 30 mN in +z-direction with various chamber pressures.

2.5. Stiffness Variations

The stiffness of cantilever beams is typically defined as the ratio of the service load (external force) to the maximum deflection. Adopting a similar definition, the stiffness of the modeled soft robot was computed on simulations results for chamber pressure ranging from $P \in [0, 15]$ kPa and for tip forces (+z-direction) ranging from $||\mathbf{F}|| \in [0, 60]$ mN. Fig. 4 depicts the variation of the soft robot's stiffness with chamber pressure and tip force. It was observed that the soft robot's stiffness changes both with the internal pressure and tip force. The

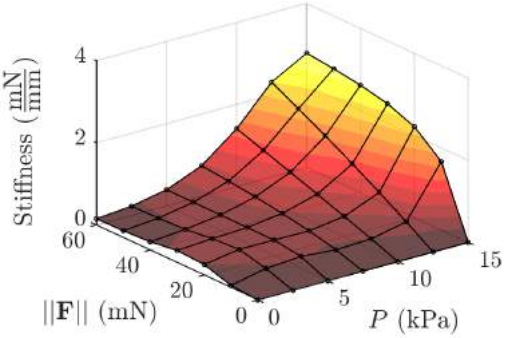


Fig. 4: Variation of the soft robot's stiffness with internal pressure and tip force.

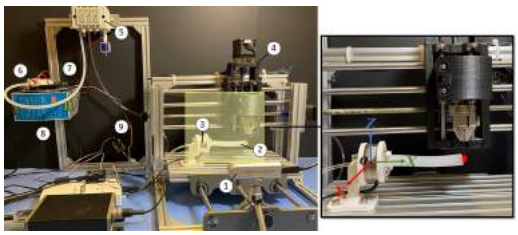


Fig. 5: Components of the mechanical and electrical modules in the prototyped soft robot (1) 3D-camera (2) soft robot (3) F/T sensor (4) 3-axis CNC machine (5) electronic pressure controller and manifold (6) air pump (7) 12-bit analog voltage generator (8) power supply (9) air pressure sensor.

variation of the stiffness with chamber pressure was deemed as the result of the longitudinal tension caused by the chamber pressure. In addition, because of the hyperelasticity incorporated to the constitutive equation (Eq. 8–9) the tip force would increase the strain thus increase the bulk moduli.

3. VALIDATION STUDY

3.1. Study Design

Four individual experiments were performed for the validation study, i.e., with chamber pressures of 0, 4, 12, and 20 kPa. With each chamber pressure, the tip of the soft robot was pulled upward from resting position while the force/torque sensor was recording the base reaction forces and torques. In parallel, a 3D camera tracked the tip position of the soft robot for validation comparison. Using the recorded force/torque data the developed model was solved and the tip computed tip positions were compared with the ground truth (recorded by the 3D camera).

3.2. Soft Robot Prototype and Experimental Setup

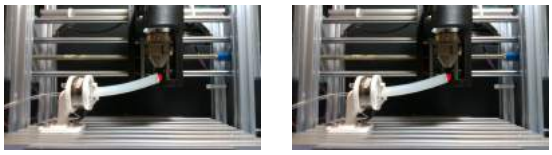
For the fabrication, a cylindrical mold was rapid-prototyped with a 3D printer (Replicator+, MakerBot, NY, USA) using PLA material.Ecoflex^{TM}00-50 (Smooth-On Inc., PA, USA) was used to make the body. Also, an platform housing was 3D-printed to install the soft robot's based on the force/torque sensor. The silicone mixture was degassed in a vaccum chamber and rested for 24 hours at $24°C$ for curing. Fig. 5 shows the experimental setup for this study. A 6-DoF force/torque sensor (ATI Industrial Automation, F/T Sensor: Mini40) was used to measure the soft robot's base reaction force and torques. An air pump (KPM27C, DC 6V, Koge Electronics) supplied the air pressure and a pressure sensor (Phidgets Inc., AB, Canada) was utilized to record the chamber's real-time pressure during the experiment. Also, an electronic pressure regulator (ITV0010-3UML, SMC, Tokyo, Japan) was used. For 3D tracking of the soft robot's tip, a 3D camera (D435i, Intel Corp., CA, USA) was used.

Table 2: Comparison of the model results with experiments.

Pressure	Tip Force	Tip Displacement		Displacement Error		Stiffness
		Model	Reference	Absolute	Relative	
(kPa)	(mN)	(mm)	(mm)	(mm)	(%)	(mN/mm)
0	51	14.4	15	0.6	3.4%	3.4
4	64	14.3	15	0.7	4.7%	4.3
12	73	13.9	15	1.1	7.3%	4.9
20	89	13.7	15	1.3	8.7%	5.9

4. RESULTS AND DISCUSSION

Fig. 6 shows representative shapes of the soft robot pulled upward 15 mm with 0 and 20 kPa chamber pressure. Also, Table 2 compares the theoretical and experimental results for the tip position for the four experiments. Similar to the simulation results, the stiffness (ratio of force to maximum deflection) It was observed that the model error increased with increasing the chamber pressure. The reason might be related to the cross-sectional expansion of the soft robot which was not neglected in the model. Nevertheless, the maximum relative error of the proposed model was 8.7%. This level of error is comparable to other studies in the literature, e.g., [13]. In addition, the post-processing showed that the stiffness of the soft robot increased from 3.4 $\frac{mN}{mm}$ ($P = 0$ kPa) to 5.9 $\frac{mN}{mm}$ ($P = 20$ kPa) indicating a 74% pressure-stiffening effect. These findings confirmed the accuracy of the proposed model to capture the effects of the internal pressure on the stiffness of the soft robots.



(a) (b)

Fig. 6: Deformed shape of the catheter with (a) 0 kPa and (b) 20 kPa chamber pressures.

5. CONCLUSION

In this study, a mechanistic model for capturing the effect of chamber pressure on the stiffness of soft pneumatic-driven robots was proposed, solved, and validated. The developed model was used to simulate the soft robot under various tip force and chamber pressure conditions. The experimental study confirmed the accuracy of the proposed model. This study was a first step toward exploiting the pressure-stiffening phenomenon for stiffness adaptation of soft surgical robots during interventional procedures. In future studies, the effects of presence of multiple chambers for directional stiffening and feasibility of position-stiffness hybrid control through tendon-pneumatic actuation will be investigated.

6. REFERENCES

[1] Mohammad Jolaei, Amir Hooshiar, Javad Dargahi, and Muthukumaran Packirisamy, "Toward task autonomy in robotic cardiac ablation: Learning-based kinematic control of soft tendon-driven catheters," *Soft Robotics*, vol. soro.2020.0006, 2020.

[2] Amir Hooshiar, Siamak Najarian, and Javad Dargahi, "Haptic telerobotic cardiovascular intervention: a review of approaches, methods, and future perspectives," *IEEE reviews in biomedical engineering*, vol. 13, pp. 32–50, 2019.

[3] Mohammad Jolaei, Amir Hooshiar, Amir Sayadi, Javad Dargahi, and Muthukumaran Packirisamy, "Sensor-free force control of tendon-driven ablation catheters through position control and contact modeling," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 5248–5251.

[4] Hesheng Wang, Runxi Zhang, Weidong Chen, Xiaozhou Wang, and Rolf Pfeifer, "A cable-driven soft robot surgical system for cardiothoracic endoscopic surgery: preclinical tests in animals," *Surgical endoscopy*, vol. 31, no. 8, pp. 3152–3158, 2017.

[5] Erina Baynojir Joyee and Yayue Pan, "Additive manufacturing of multi-material soft robot for on-demand drug delivery applications," *Journal of Manufacturing Processes*, vol. 56, pp. 1178–1184, 2020.

[6] Olatunji Mumini Omisore, Shipeng Han, Jing Xiong, Hui Li, Zheng Li, and Lei Wang, "A review on flexible robotic systems for minimally invasive surgery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.

[7] Mahta Khoshnam, Allan C Skanes, and Rajni V Patel, "Modeling and estimation of tip contact force for steerable ablation catheters," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 5, pp. 1404–1415, 2015.

[8] Vincent A Aloi and D Caleb Rucker, "Estimating loads along elastic rods," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 2867–2873.

[9] Shahir Hasanzadeh and Farrokh Janabi-Sharifi, "Model-based force estimation for intracardiac catheters," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 1, pp. 154–162, 2015.

[10] Robert K Katzschmann, Cosimo Della Santina, Yasunori Toshimitsu, Antonio Bicchi, and Daniela Rus, "Dynamic motion control of multi-segment soft robots using piecewise constant curvature matched with an augmented rigid body model," in *2019 2nd IEEE International Conference on Soft Robotics (RoboSoft)*. IEEE, 2019, pp. 454–461.

[11] Robert J Webster III and Bryan A Jones, "Design and kinematic modeling of constant curvature continuum robots: A review," *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1661–1683, 2010.

[12] Amir Hooshiar, Amir Sayadi, Mohammad Jolaei, and Javad Dargahi, "Accurate estimation of tip force on tendon-driven catheters using inverse cosserat rod model," in *2020 International Conference on Biomedical Innovations and Applications (BIA)*. IEEE, 2020, pp. 37–40.

[13] John Till, Vincent Aloi, and Caleb Rucker, "Real-time dynamics of soft and continuum robots based on cosserat rod models," *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 723–746, 2019.

[14] Richard M Murray, Zexiang Li, S Shankar Sastry, and S Shankara Sastry, *A mathematical introduction to robotic manipulation*, CRC press, 1994.

INFORMATION-BOTTLENECK-BASED BEHAVIOR REPRESENTATION LEARNING FOR MULTI-AGENT REINFORCEMENT LEARNING

Yue Jin[1] *Shuangqing Wei*[2] *Jian Yuan*[1] *Xudong Zhang*[1]

[1] Department of Electronic Engineering, Tsinghua University, Beijing, China
[2] School of Electrical Engineering and Computer Science, Louisiana State University, Baton Rouge, USA

ABSTRACT

In multi-agent deep reinforcement learning, extracting sufficient and compact information of other agents is critical to attain efficient convergence and scalability of an algorithm. In canonical frameworks, distilling of such information is often done in an implicit and uninterpretable manner, or explicitly with cost functions not able to reflect the relationship between information compression and utility in representation. In this paper, we present Information-Bottleneck-based Other agents' behavior Representation learning for Multi-agent reinforcement learning (IBORM) to explicitly seek low-dimensional mapping encoder through which a compact and informative representation relevant to other agents' behaviors is established. IBORM leverages the information bottleneck principle to compress observation information, while retaining sufficient information relevant to other agents' behaviors used for cooperation decision. Empirical results have demonstrated that IBORM delivers the fastest convergence rate and the best performance of the learned policies, as compared with implicit behavior representation learning and explicit behavior representation learning without explicitly considering information compression and utility.

Index Terms— Multi-agent deep reinforcement learning, representation learning, information bottleneck principle

1. INTRODUCTION

Representation learning, which aims to learn informative and effective features of a task, is a key part of deep learning. Naturally, deep reinforcement learning (DRL) is expected to benefit from the help of representation learning. Many works [1–4] have dedicated to boost DRL by learning a compact, discriminative or task-relevant representation from observations. However, in multi-agent tasks, a good task-relevant representation also needs to be teammate-relevant or opponent-relevant. It has been demonstrated in some works [5–7] that using independent DRL (ignore other agents' behaviors) may lead to unsatisfactory results. Meanwhile, some works [8, 9] indicate that inferring other agents' policies can improve cooperation between agents, but is prohibitively expensive for policies parameterized by deep neural networks. These studies imply the demand for more efficient and effective representation learning of other agents in multi-agent DRL (MADRL).

Two core problems of representing other agents in MADRL are what to represent and how to combine the representation with MADRL. Foerster et al. [7] leverage low-dimensional fingerprints to represent other agents' policy changes, which forms succinct features, but is deficient in policy information completeness. Jin et al. [6] propose to represent other agents' behaviors implicitly using their positions at adjacent timesteps. However, the behavior representation is learned via MADRL in an implicit and uninterpretable manner. He et al. [5] propose deep reinforcement opponent network (DRON) to learn representations of other agents' actions explicitly by leveraging other agents' actions as supervision signals. However, the compactness and information utility of representation are not considered.

In this paper, we leverage information bottleneck principle [10, 11] to learn an informative and compact representation relevant to other agents' behaviors to improve the performance of MADRL. In particular, we employ an encoder to extract features from each agent's positions at two adjacent timesteps, based on which a classifier is learned to estimate actions of each agent. To filter out irrelevant information from observations and retain sufficient amount of information of other agents' actions used for cooperation decision, we follow the information bottleneck principle to minimize the mutual information between the representation and the observations, while maximizing the mutual information between the representation and other agents' actions. To this end, we adopt a variational method [12] to estimate the two mutual information and integrate this process into behavior representation learning. We combine our proposed behavior representation learning method with our recent work, stabilized multi-agent deep Q learning (SMADQN) [6] by multi-task learning and thereby the learned representation can also retain other information about the MADRL task in addition to other agents' actions. Experimental results demonstrate the superior performance of our proposed method compared to vanilla SMADQN and DRON-based SMADQN.

This work was supported in part by the National Natural Science Foundation of China under Grant U20B2060.

In summary, the main contributions of this paper are as follows:

1) We propose an information-bottleneck-based behavior representation learning method through which compact and informative features of other agents' behaviors are learned and exploited to facilitate MADRL.

2) We conduct extensive experiments in cooperative navigation tasks [6, 8, 13]. Experimental results demonstrate that compared to implicit behavior representation learning and the explicit behavior representation learning that does not consider information utility and compression, our method performs best in terms of both learning speed and the success rates of the resulting policies.

2. METHOD

In this section, we first introduce the Markov game, SMADQN [6] and DRON [5]. Then, we present our method.

A Markov game with N agents involves a set of states s, joint actions $(a_1, \cdots, a_N)$, transition probability function $p(s'|s, a_1, \cdots, a_N)$, and each agent's reward function $r_i(s, a_1, \cdots, a_N)$, $i \in [1, N]$. At each timestep, each agent executes an action according to its policy π_i. A problem of Markov game is to find the optimal policy π_i^* for each agent so that $\forall \pi_i, R_i(s^t, \pi_1^*, \cdots, \pi_i^*, \cdots, \pi_N^*) \geq R_i(s^t, \pi_1^*, \cdots, \pi_i, \cdots, \pi_N^*)$, where $R_i(s^t, \pi_1, \cdots, \pi_N) = E[\sum_{\tau=t}^{T} \gamma^{\tau-t} r_i(s^\tau, a_1^\tau, \cdots, a_N^\tau)]$ denotes the expected total reward of agent i, T is the time horizon, $\gamma \in [0, 1]$ is a discount factor. For convenience, we use r_i^t to denote the reward of agent i at timestep t.

SMADQN defines an extended action-value function G for each agent to measure its expected total reward when it follows policy π_i. For agent i, G-function is defined as

$$G_i^{\pi_i}(s^t, s_{-i}^t, s_{-i}^{t+1}, a_i^t) = Q_i^{\pi_i}(s^t, f(s_{-i}^t, s_{-i}^{t+1}), a_i^t), \tag{1}$$

where s^t represents global states, s_{-i}^t and s_{-i}^{t+1} represent states of agents except agent i at two adjacent timesteps, f is an action estimation function of other agents' actions. G-function is a composite function that incorporates the action estimation function into the original action-value function Q [14]. An approximate extended Bellman equation for the optimal G-function is derived as:

$$\begin{aligned} &\mathbb{E}_{s_{-i}^{t+1}|s_{-i}^t, a_{-i}^t} G_i^*(s^t, s_{-i}^t, s_{-i}^{t+1}, a_i^t) \approx \\ &\mathbb{E}_{s^{t+1}|s^t, a_i^t, a_{-i}^t} \left[r_i^{t+1} + \gamma \max_{a_i^{t+1}} G_i^*(s^{t+1}, s_{-i}^t, s_{-i}^{t+1}, a_i^{t+1}) \right]. \end{aligned} \tag{2}$$

The optimal G-function is approximated by a neural network learned by minimizing the loss function given as:

$$\begin{aligned} L = \mathbb{E}_{s^t, s^{t+1}, a_i^t} \Bigg[\Bigg(& r_i^{t+1} + \gamma \max_{a_i^{t+1}} G_i(s^{t+1}, s_{-i}^t, s_{-i}^{t+1}, a_i^{t+1}) \\ & - G_i(s^t, s_{-i}^t . s_{-i}^{t+1}, a_i^t) \Bigg)^2 \Bigg], \end{aligned} \tag{3}$$

SMADQN learns action estimation function of other agents' actions implicitly, which may lead to trivial estimation performance and thereby cause limited performance of the resulting policies.

Instead of merging action estimation learning into MADRL, DRON leverages other agents' actions as supervision signals and adopts supervised learning to learn action estimation. It uses a classification network to estimate other agents' actions. The output of the last hidden layer of the network is used as other agents' action representation, and is fed into a decision network. DRON can learn representations of other agents' behaviors explicitly. However, it does not consider information compression and retention in the representation.

To facilitate and improve MADRL, extracting informative and compressed representation of other agents' behaviors is critical. To this end, we propose Information-Bottleneck-based Other agents' behavior Representation learning for Multi-agent reinforcement learning (IBORM), which equips a behavior representation with the following capabilities, a) to extract features of other agents' actions, b) to filter out irrelevant information while retaining sufficient information about the actions and other potentially helpful information about the task to facilitate MADRL. Specifically, we implement our idea in SMADQN. We replace the implicit action representation learning of SMADQN with explicit action representation learning. An encoder is employed to learn the representation using other agents' states at adjacent timesteps as inputs. The encoder's output is a low-dimensional feature vector of other agents' actions, from which a classifier can predict the actions. The representation learning and SMADQN are combined by leveraging multi-task learning.

Additionally, to learn an informative and compressed representation, we leverage information bottleneck principle [10, 11] to constrain the information contained in the representation. To be specific, information bottleneck (IB) principle introduces an information theory principle for extracting an optimal representation Z that captures the relevant information in a random variable X about another correlated random variable Y while minimizing the amount of irrelevant information, where (Y, X, Z) forms a Markov chain, $Y \rightarrow X \rightarrow Z$. Namely, finding the optimal representation function is formulated as minimizing the following Lagrangian

$$\mathcal{L}(p(z|x)) = I(X; Z) - \kappa I(Z; Y), \tag{4}$$

where κ determines how much relevant information is contained in the representation. Based on IB principle, we constrain the representation learning by the following terms

$$\mathcal{L}(\alpha) \triangleq I(\phi_{s_{-i,j}}; ENC_i^\alpha(\phi_{s_{-i,j}}) - \kappa I(ENC_i^\alpha(\phi_{s_{-i,j}}); a_{-i,j}), \tag{5}$$

where α denotes the parameters of the encoder, $a_{-i,j}$ and $\phi_{s_{-i,j}}$ denote the action and agent i's observation of the jth agent other than agent i, respectively.

Overall, the loss function of IBORM is given as:

$$L_i(\alpha, \beta, \theta) = J_i^{CE}(\alpha, \beta) + \lambda_1 J_i^{DRL}(\alpha, \theta) + \lambda_2 \mathcal{L}(\alpha), \tag{6}$$

where $J_i^{CE}(\alpha, \beta)$ denotes the cross-entropy between classifier's output and each agent's true action , α and β are parameters of the encoder and the classifier. $J_i^{DRL}(\alpha, \theta)$ denotes a

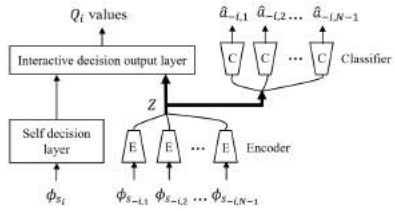


Fig. 1: Network architecture diagram of IBORM.

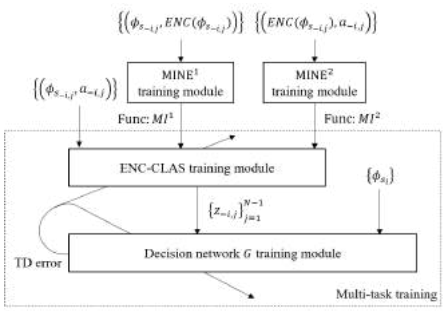


Fig. 2: Framework of IBORM algorithm.

modified loss function of SMADQN, defined as

$$J_i^{DRL}(\alpha,\theta)=\mathbb{E}_{\phi_{s_i}^t,a_i^t,\phi_{s-i,1}^{t+1},\cdots,\phi_{s-i,N-1}^{t+1}} \left[\left(y_i^t-G_i^\theta(\phi_{s_i}^t,ENC_i^\alpha(\phi_{s-i,1}^{t+1}),\cdots,ENC_i^\alpha(\phi_{s-i,N-1}^{t+1}),a_i^t)\right)^2\right], \tag{7}$$

where $y_i^t=r_i^{t+1}+\gamma\max_{a_i}G_i^{\theta_{tar}}(\phi_{s_i}^{t+1},ENC_i^\alpha(\phi_{s-i,1}^{t+1}),\cdots,ENC_i^\alpha(\phi_{s-i,N-1}^{t+1}),a_i)$. θ and θ_{tar} are parameters of G network and target network [15], respectively. Compared with (3), we replace other agents' adjacent states with $N-1$ behavior representations. For notation convenience, we rewrite (6) as

$$L_i(\alpha,\beta,\theta)=J_i^{CE}(\alpha,\beta)+\lambda_1 J_i^{DRL}(\alpha,\theta)+ \lambda_2 I(\phi_{s_{-i,j}},ENC_i^\alpha(\phi_{s_{-i,j}}))-\lambda_3 I(ENC_i^\alpha(\phi_{s_{-i,j}}),a_{-i,j}), \tag{8}$$

where $\lambda_1,\lambda_2,\lambda_3$ are positive weights. From the perspective of information utility, the first and the last terms of (8) are for extracting sufficient information of other agents' behaviors. The second term extracts relevant information of the task. The third term filters out irrelevant information. Compared to IBORM, SMADQN only uses the DRL-based term, where the behavior representation learning is implicitly contained. DRON uses the cross-entropy term but does not constrain the amount and utility of the information in the representation.

The network architecture of IBORM is shown in Fig. 1. An encoder-classifier is used to estimate each of other agents' actions. The encoder is duplicated by $N-1$ times and generates bottleneck representations for $N-1$ other agents' actions, respectively. Then, the representations are incorporated into a decision network to make interactive decisions.

To estimate the two mutual information terms in (8), we employ the Mutual Information Neural Estimator (MINE) [12] that estimates the mutual information between two variables X and Z as $\widehat{I(X,Z)}=\sup_{\omega\in\Omega}\mathbb{E}_{\mathbb{P}_{XZ}}[T_\omega(x,z)]-\log(\mathbb{E}_{\mathbb{P}_X\otimes\mathbb{P}_Z}[e^{T_\omega(x,z)}])$ by leveraging a trainable neural network

Algorithm 1 Stabilized multi-agent deep Q-learning with information-bottleneck-based other agents' behavior representation learning

for agent $i=1$ to N **do**
  Initialize networks $ENC_i^\alpha:\phi_{s_{-i,j}}\to z_{-i,j}, CLAS_i^\beta:z_{-i,j}\to\hat{a}_{-i,j}$, $MINE_i^{\omega_1}(\phi_{s_{-i,j}},z_{-i,j}), MINE_i^{\omega_2}(z_{-i,j},a_{-i,j})$, $G_i^\theta(\phi_{s_i},ENC_i^\alpha(\phi_{s_{-i,1}}),\cdots,ENC_i^\alpha(\phi_{s_{-i,N-1}}),a_i)$, target network $G_i^{\theta_{tar}}$ with $\theta_{tar}\leftarrow\theta$, replay buffer $\mathcal{D}_i$
end for
for episode$=1$ to Z **do**
  Receive $\phi_{s_i}^1,\boldsymbol{\phi}_{s_{-i}}^1$ for each agent
  for t$=1$ to T **do**
    Execute action for each agent i: $a_i^t=\arg\max_{a_i}G_i^\theta(\phi_{s_i}^t,ENC_i^\alpha(\phi_{s_{-i,1}}^t),\cdots,ENC_i^\alpha(\phi_{s_{-i,N-1}}^t),a_i)$
    Receive $r_i^{t+1},\phi_{s_i}^{t+1},\boldsymbol{\phi}_{s_{-i}}^{t+1}$ for each agent i
    Record data: $\mathcal{D}_i\leftarrow\mathcal{D}_i\bigcup\{(\phi_{s_i}^t,a_i^t,r_i^{t+1},\phi_{s_i}^{t+1},\boldsymbol{\phi}_{s_{-i}}^{t+1},\boldsymbol{a}_{-i}^t)\}$ for each agent i
    for agent $i=1$ to N **do**
      Sample M tuples $\{(\phi_{s_i}^{\tau_k},a_i^{\tau_k},r_i^{\tau_k+1},\phi_{s_i}^{\tau_k+1},\boldsymbol{\phi}_{s_{-i}}^{\tau_k+1},\boldsymbol{a}_{-i}^{\tau_k})\}_{k=1}^M$ from $\mathcal{D}_i$
      Compute $\{z_{-i,j}^{\tau_k+1}=ENC_i^\alpha(\phi_{s_{-i,j}}^{\tau_k+1})\}_{j=1}^{N-1}$
      Compute $y_i^{\tau_k}$ by $y_i^{\tau_k}=r_i^{\tau_k+1}+\gamma\max_{a_i}G_i^{\theta_{tar}}(\phi_{s_i}^{\tau_k+1},z_{-i,1}^{\tau_k+1},\cdots,z_{-i,N-1}^{\tau_k+1},a_i)$
      Update ω_1,ω_2 according to [12] using SGD
      Update α,β,θ to minimize (8) using SGD
      Update target network with soft update rate η: $\theta_{tar}\leftarrow\eta\theta+(1-\eta)\theta^{tar}$
    end for
  end for
end for

T_ω with parameters ω. Specifically, we use two networks corresponding to MINEs of the two mutual information terms in (8). To integrate learning of MINEs and IBORM, we adopt an interlaced learning manner to update parameters of MINEs and IBORM alternately. A framework of our algorithm is shown in Fig. 2. The complete algorithm is shown in Algorithm 1, where we denote $\boldsymbol{\phi}_{s_{-i}}^t=[\phi_{s_{-i,1}}^t,\cdots,\phi_{s_{-i,N-1}}^t]$ and $\boldsymbol{a}_{-i}^t=[a_{-i,1}^t,\cdots,a_{-i,N-1}^t]$ for notation convenience.

3. EXPERIMENTS

In this section, we evaluate IBORM in multi-agent cooperative navigation task with the same settings used in [6]. In this task, agents need to cooperate through motions to reach the same number of targets using the minimum time. An example containing three agents and targets is illustrated in Fig. 3. At each timestep, each agent selects a target and move a fixed distance toward the target. The action of an agent is defined as $a_i\in[1,N]$ that indicates the index of the target selected by it. Agents' speed is 1 m/timestep. The current observation of agent i, i.e. ϕ_{s_i}, is composed of the current positions of other entities (targets and the other agents) and agent i's last action. An agent's observation about other agents' states, i.e. $\phi_{s_{-i}}$, includes the positions of targets and the current and last positions of other agents, which are necessary for an agent to predict other agents' actions. The size of the environment is

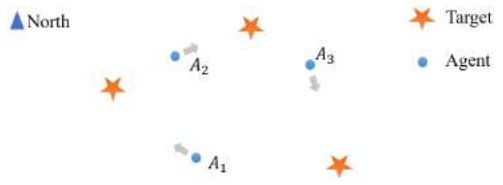


Fig. 3: Illustration of the cooperative navigation task involving three agents.

Table 1: Test results of different methods.

Method	Success rate				
	N=3	N=4	N=5	N=6	N=7
SMADQN	98.2%	97.8%	96.1%	91.2%	0.0%
SMADRON	98.9%	96.9%	92.9%	93.5%	82.5%
IBORM (ours)	**99.3%**	**98.1%**	**97.1%**	**93.5%**	**87.8%**

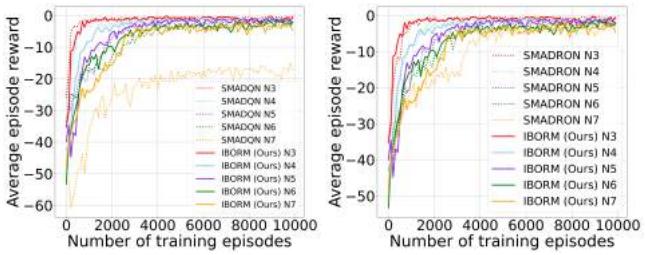
$15 \times 15\ m^2$. The maximum episode length is 30 timesteps. Agents are homogeneous. They share a common policy and reward function. The reward function is aligned with [6].

Network structure and hyperparameters of IBORM are as follows. The decision network has two hidden layers containing 300 and 200 units, respectively. The encoder and classifier both have two hidden layers. Encoder has 32 and 16 units in its two layers, respectively. Classifier has 16 and 32 units in its two layers, respectively. Adam optimizer with a learning rate of 0.01 is applied to update parameters. The size of the replay buffer is 1500, the batch size of SGD is 32, $\gamma = 1$ and the target network is updated with a soft update rate [16] $\eta = 0.001$, which are all aligned with SMADQN. λ_1, λ_2 and λ_3 are set as 10, 0.001 and 0.1 by grid search method.

We compare IBORM with SMADQN and a comparable DRON whose DRL loss are aligned with SMADQN. We name the latter SMADRON. For SMADQN, its network has two hidden layers containing the the same number of units as that in IBORM's decision network. For SMADRON, its network structure is the same as IBORM. Hyperparameters used in SMADQN and SMADRON are the same as IBORM.

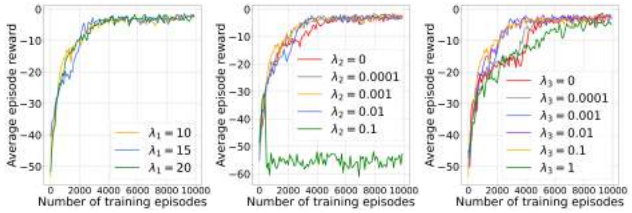
We train each method by 10k episodes. At the beginning of each episode, positions of targets and agents are generated randomly. Each method is evaluated with different numbers of targets and agents ($N = 3, 4, 5, 6, 7$). Convergence curves of average episode reward are shown in Fig.4. As we can see from the results, when the number of agents increases, IBORM learns faster than the other two methods, which indicates the advantage of IBORM over implicit behavior representation learning (SMADQN) and explicit behavior representation learning without considering information utility (SMADRON). To test the performance of the learned policies, we generate 1000 testing tasks with random positions of targets and agents. Table 1 shows success rate of cooperative navigation with different policies, from which we can see that IBORM outperforms the other two methods consistently.

We also investigate how learning performance of IBORM changes when λ_1, λ_2, and λ_3 vary in (8). Fig.5 shows the results in tasks of $N = 6$. Three subfigures corresponds to λ_1, λ_2 and λ_3, respectively. As shown in the results, the suitable value range of each λ is relatively wide. Fig.5(b) indicates large λ_2 can lead to failure of IBORM. This is because large λ_2 causes much information to be discarded. Fig.5(c) shows that large λ_3 reduces learning speed and gets less rewards, because when λ_3 is enlarged, λ_1 is weakened relatively and thus the learned representation captures insufficient information regarding the task. Too small λ_3 also slows down learning, because the learned representation captures insufficient information about other agents' behaviors and thus provides less help to MADRL. Additionally, when λ_2 or λ_3 equals zero, the learning performance degenerates, which indicates the importance of each mutual information constraint in IBORM.



(a) (b)

Fig. 4: Convergence curves of average episode reward of different methods. (a) IBORM vs. SMADQN. (b) IBORM vs. SMADRON.



(a) (b) (c)

Fig. 5: Sensitivity of IBORM to λ values on tasks containing six targets and agents. Subfigures (a), (b) and (c) corresponds to λ_1, λ_2 and λ_3, respectively.

4. CONCLUSION

We propose IBORM to facilitate MADRL by learning a compact and informative representation regarding other agents' behaviors. We implement IBORM based on our recently proposed MADRL algorithm, SMADQN, by replacing the implicit behavior representation learning of SMADQN with information-bottleneck-based explicit behavior representation learning. Experimental results demonstrate that IBORM learns faster and the resulting policies can achieve higher success rate consistently, as compared with implicit behavior representation learning (SMADQN) and explicit behavior representation learning (SMADRON) without considering information compression and utility.

5. REFERENCES

[1] S. Lange and M. Riedmiller, "Deep auto-encoder neural networks in reinforcement learning," in *Proceedings of IEEE International Joint Conference on Neural Networks*, 2010, pp. 1–8.

[2] M. Laskin, A. Srinivas, and P. Abbeel, "Curl: Contrastive unsupervised representations for reinforcement learning," in *Proceedings of International Conference on Machine Learning*, 2020, pp. 5639–5650.

[3] V. Pacelli and A. Majumdar, "Learning task-driven control policies via information bottlenecks," *arXiv preprint arXiv:2002.01428*, 2020.

[4] V. François-Lavet, Y. Bengio, D. Precup, and J. Pineau, "Combined reinforcement learning via abstract representations," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 3582–3589.

[5] H. He, J. Boyd-Graber, K. Kwok, and H. Daumé III, "Opponent modeling in deep reinforcement learning," in *Proceedings of International Conference on Machine Learning*, 2016, pp. 1804–1813.

[6] Y. Jin, S. Wei, J. Yuan, X. Zhang, and C. Wang, "Stabilizing multi-agent deep reinforcement learning by implicitly estimating other agents' behaviors," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020, pp. 3547–3551.

[7] J. Foerster, N. Nardelli, G. Farquhar, T. Afouras, P. H. Torr, P. Kohli, and S. Whiteson, "Stabilising experience replay for deep multi-agent reinforcement learning," in *Proceedings of International Conference on Machine Learning*, 2017, pp. 1146–1155.

[8] R. Lowe, Y. Wu, A. Tamar, J. Harb, O. P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *Advances in Neural Information Processing Systems*, 2017, pp. 6379–6390.

[9] G. Tesauro, "Extending q-learning to general adaptive multi-agent systems," in *Advances in Neural Information Processing Systems*, 2004, pp. 871–878.

[10] N. Tishby and N. Zaslavsky, "Deep learning and the information bottleneck principle," in *IEEE Information Theory Workshop*, 2015, pp. 1–5.

[11] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," *arXiv preprint physics/0004057*, 2000.

[12] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, "Mutual information neural estimation," in *Proceedings of International Conference on Machine Learning*, 2018, pp. 531–540.

[13] Y. Jin, Y. Zhang, J. Yuan, and X. Zhang, "Efficient multi-agent cooperative navigation in unknown environments with interlaced deep reinforcement learning," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2019, pp. 2897–2901.

[14] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[15] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[16] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," in *Proceedings of International Conference on Learning Representations*, 2016.

ESTIMATION OF FIELDS USING BINARY MEASUREMENTS FROM A MOBILE AGENT

Alex S. Leong and Mohammad Zamani

Land Division, Defence Science and Technology Group, Australia
{alex.leong, mohammad.zamani}@dst.defence.gov.au

ABSTRACT

In this paper we consider the problem of field estimation using a mobile autonomous agent. The agent is equipped with a sensor which can make noisy binary measurements of the field. We model the field as a sum of radial basis functions, whose parameters are then estimated using sequential Monte Carlo (SMC) techniques. We also devise an active sensing mechanism for the agent to adaptively choose its next measurement location, given the information currently collected. Simulation studies illustrate the performance of the proposed algorithms.

1. INTRODUCTION

The occurrence of a CBRN (chemical, biological, radiological, or nuclear) event can pose substantial risk to health and hinder efforts to perform operations (humanitarian / disaster relief / military) within the vicinity. In order to enable operations within contaminated areas, a crucial task is to map out such areas accurately and efficiently. We consider in this paper the field estimation problem using a mobile autonomous agent, where we try to estimate (or construct a model which can provide estimates of) the field values / contamination levels at each possible location in the affected area.

Models on gas and chemical dispersion include various statistical models [1], while models for radiological sources include Poisson distributions [2]. In this paper, rather than assuming specific characteristics of a particular type of threat, we will model the field as a sum of radial basis functions, as has also been considered in e.g. [3–5]. Estimation of the field then reduces to estimation of the parameters of the individual basis functions, which in this paper will be carried out using sequential Monte Carlo (SMC) techniques [6].

In environment monitoring using sensor networks, many works assume that a large number of sensors at (mostly) fixed locations can make measurements in order to map out an environment. For CBRN events, since we often won't know specifically when and where they will occur, such an assumption may not necessarily be realistic. We instead consider a setup with a mobile robot or autonomous vehicle, which we will also refer to as an *agent*, that can move around to explore and construct a map of the contaminated environment. We aim to actively choose the measurement locations, based on the information collected up to that point, which we call *active sensing* [7]. For radiological sources, measurements are usually assumed to be fairly accurate, see e.g. [2], [8]. In the current work, we consider coarsely quantized measurements, in particular binary measurements, which may be a more appropriate model for chemical sensors [9], [10].

Previous works on field estimation using mobile agents include [3, 5], with noisy but non-binary measurements. Binary measurements are considered for field monitoring in [11], but with sensors randomly deployed at fixed locations. The use of mobile agents with binary measurements for source localization is considered in [12], however only a single source is assumed. Localization of radiological sources using mobile agents is studied in [13], where the measurements are non-binary.

The paper is organized as follows. The system model and problem statement are presented in Section 2. The SMC approach to parameter estimation is presented in Section 3, with active sensing considered in Section 4. Simulation studies are given in Section 5.

2. SYSTEM MODEL AND PROBLEM STATEMENT

The model of the field is given by

$$\phi(\mathbf{x}) = \sum_{j=1}^{J} \beta_j K_j(\mathbf{x}) \tag{1}$$

where $\phi(\mathbf{x})$ is the value of the field at position $\mathbf{x} \in \mathbb{R}^2$, $K_j(.) : \mathbb{R}^2 \to \mathbb{R}$ are basis functions, and $\beta_j \in \mathbb{R}$ are coefficients multiplying the basis functions. Similar models have been used in e.g. [3–5]. For the basis functions, in this paper we will use the Gaussian kernel

$$K_j(\mathbf{x}) = \exp\left(-\frac{||\mathbf{c}_j - \mathbf{x}||^2}{\sigma_j^2}\right), \tag{2}$$

where $\mathbf{c}_j \in \mathbb{R}^2$ and $\sigma_j \in \mathbb{R}$ represent the "center" and the "width" of the j-th basis function, respectively. The Gaussian kernel is an example of a radial basis function. It is known that the class of radial basis functions is sufficiently rich to allow many fields to be approximated to arbitrary accuracy (as the number of basis functions $J \to \infty$) [14], see also [4, 5].

For the measurement model, at position $\mathbf{x}$, the sensor on-board the agent is first assumed to have a noisy reading $y(\mathbf{x}) = \phi(\mathbf{x}) + v(\mathbf{x})$, where $v(\mathbf{x}) \sim \mathcal{N}(0, \sigma_v^2)$ is i.i.d. Gaussian noise. In this paper we will assume that the noise variance σ_v^2 is unknown, since this may depend on factors such as the type of contaminant and the particular sensor technology. This noisy reading is then thresholded to give a binary measurement

$$z(\mathbf{x}) = \mathbb{1}(y(\mathbf{x}) > \tau),$$

where $\mathbb{1}(.)$ is the indicator function and τ is the known threshold.

We assume that $\mathbf{c}_j, \sigma_j^2, j = 1, \ldots, J$ can be chosen by us. The objective is to estimate the field (via estimation of the field parameters $\beta_j, j = 1, \ldots, J$) from binary measurements collected by an agent, where the agent can collect measurements from any position $\mathbf{x}$ in an area of interest $\mathcal{S}$.

Preliminaries: If σ_v^2 and $\beta_j, j = 1, \ldots, J$ are known, then the

probabilities of receiving a 0 or 1 at location $\mathbf{x}$ are:

$$\begin{aligned}\mathbb{P}(z(\mathbf{x})=0) &= \mathbb{P}(y(\mathbf{x})<\tau)\\ &= \mathbb{P}\Big(v(\mathbf{x})<\tau-\sum_{j=1}^{J}\beta_j\exp\Big(-\frac{||\mathbf{c}_j-\mathbf{x}||^2}{\sigma_j^2}\Big)\Big)\\ &= \Phi\Big(\frac{1}{\sigma_v}\Big(\tau-\sum_{j=1}^{J}\beta_j\exp\Big(-\frac{||\mathbf{c}_j-\mathbf{x}||^2}{\sigma_j^2}\Big)\Big)\Big)\\ \mathbb{P}(z(\mathbf{x})=1) &= 1-\mathbb{P}(z(\mathbf{x})=0), \end{aligned} \quad (3)$$

where $\Phi(x)\triangleq\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt$ is the cumulative distribution function (cdf) of the standard normal distribution $\mathcal{N}(0,1)$.

3. SMC FOR PARAMETER ESTIMATION

As mentioned in the previous section, we assume that $\mathbf{c}_j,\sigma_j^2, j=1,\ldots,J$ can be chosen/fixed at certain values, while $\beta_j, j=1,\ldots,J$ are to be estimated. In our proposed SMC approach, the unknown σ_v^2 will also be estimated. Define the parameter vector:

$$\triangleq(\beta_1,\ldots,\beta_J,\log\sigma_v). \quad (4)$$

We use $\log\sigma_v$ rather than σ_v^2 in (4) as we want to ensure that our estimates of σ_v^2 remain positive even when generating new particles (of $\log\sigma_v$) from a Gaussian distribution using the approach of [15], i.e. that estimates of $\sigma_v^2=\exp(2\log\sigma_v)>0$ even when $\log\sigma_v<0$. On the other hand, we allow β_j to possibly take on negative values, as the approximation theory results of [14] require that $\beta_j\in\mathbb{R}$. Note however that depending on the application, the field values $\phi(\mathbf{x})$ may be strictly non-negative. In our estimates for the fields in Section 5, to enforce non-negativity we will use

$$\hat{\phi}(\mathbf{x})=\max\left\{0,\sum_{j=1}^{J}\hat{\beta}_j\exp\Big(-\frac{||\hat{\mathbf{c}}_j-\mathbf{x}||^2}{\hat{\sigma}_j^2}\Big)\right\}. \quad (5)$$

We wish to compute the posterior probability density function (pdf)[1] $p(\ |z_{1:k};\mathbf{x}_{1:k})$, where

$$z_{1:k}\triangleq\{z_1,z_2,\ldots,z_k\},\quad \mathbf{x}_{1:k}\triangleq\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}.$$

From the posterior pdfs, one can compute different types of estimates of such as conditional mean estimates.

In general, computation of posterior pdfs cannot be done analytically and must be approximated. If one wants to estimate only after an entire set of measurements have been collected, then the posterior pdf can be approximated using, e.g., Markov Chain Monte Carlo (MCMC) methods [18]. In this paper however, we are interested in computing $p(\ |z_{1:k};\mathbf{x}_{1:k})$ at every k, and then using the updated posterior to inform our choice of the next measurement location $\mathbf{x}_{k+1}$, see Section 4. Thus sequential Monte Carlo (SMC) methods [6] are more appropriate for our situation.

We first present in Algorithm 1 an algorithm for parameter estimation using SMC, which is based on [15]. The scheme of [15] in turn is based on the auxiliary particle filter [19], but modified to handle estimation of static parameters rather than time-varying states. Algorithm 1 approximates the posterior pdf $p(\ |z_{1:k};\mathbf{x}_{1:k})$ by a set of particles $^{(i)}_{k}, i=1,\ldots,N$, and associated weights $\boldsymbol{w}_k^{(i)}, i=1,\ldots,N$. From this approximation conditional mean estimates at iteration k can be computed as:

$$\hat{\beta}_{j,k}=\sum_{i=1}^{N}\boldsymbol{w}_k^{(i)}\theta_{j,k}^{(i)},\ j=1,\ldots,J,\ \hat{\sigma}_{v,k}^2=\sum_{i=1}^{N}\boldsymbol{w}_k^{(i)}\exp(2\theta_{J+1,k}^{(i)})$$

where we denote $^{(i)}_{k}\triangleq(\theta_{1,k}^{(i)},\ldots,\theta_{J+1,k}^{(i)})$. In lines 9 and 15 of Algorithm 1, the likelihood functions can be computed in a similar manner to (3) as:

$$\begin{aligned}&p(z_k=0|\ ^{(i)}_{k};\mathbf{x}_k)\\ &=\Phi\left(\frac{1}{\exp(\theta_{J+1,k}^{(i)})}\left(\tau-\sum_{j=1}^{J}\theta_{j,k}^{(i)}\exp\left(-\frac{||\mathbf{c}_j-\mathbf{x}_k||^2}{\sigma_j^2}\right)\right)\right)\\ &p(z_k=1|\ ^{(i)}_{k};\mathbf{x}_k)=1-p(z_k=0|\ ^{(i)}_{k};\mathbf{x}_k),\end{aligned} \quad (6)$$

and similarly for $p(z_k|\mathbf{m}_k^{(i)};\mathbf{x}_k)$ and $p(z_k|\mathbf{m}_k^{(i^-)};\mathbf{x}_k)$.

In line 14 of Algorithm 1, we have introduced a (small) parameter $\eta\geq 0$, so that the variance of the distribution which we sample particles from is $h^{2-\eta}\mathbf{V}_{k-1}$, which is slightly larger than the variance $h^2\mathbf{V}_{k-1}$ used in [15]. We found that this alleviates somewhat the problem of parameter estimates becoming stuck at possibly inaccurate values and unable to recover, although it also increases the variance in the estimates.

Algorithm 1 SMC algorithm for parameter estimation

1: **Algorithm Parameters**: $N\in\mathbb{N}$, $a\in(0,1)$, $h=\sqrt{1-a^2}$, $\eta\geq 0$, prior pdf $p_0(\ )$
2: **Inputs**: Measurement locations $\{\mathbf{x}_k\}$
3: **Outputs**: Particles $\{\ ^{(i)}_{k}\}$ and weights $\{\boldsymbol{w}_k^{(i)}\}$
4: Sample particles $^{(i)}_{0}, i=1,\ldots,N$ from $p_0(\ )$, and assign weights $\boldsymbol{w}_0^{(i)}=\frac{1}{N}, i=1,\ldots,N$
5: **for** $k=1,2,\ldots,$ **do**
6: Observe z_k at location $\mathbf{x}_k$
7: **for** $i=1,\ldots,N$ **do**
8: Compute $\mathbf{m}_{k-1}^{(i)}=a\ ^{(i)}_{k-1}+(1-a)\ ^{-}_{k-1}$, where $^{-}_{k-1}=\sum_{i=1}^{N}\boldsymbol{w}_{k-1}^{(i)}\ ^{(i)}_{k-1}$
9: Assign $\tilde{\boldsymbol{w}}_k^{(i)}\propto p(z_k|\mathbf{m}_{k-1}^{(i)};\mathbf{x}_k)\boldsymbol{w}_{k-1}^{(i)}$
10: **end for**
11: Normalize $\{\tilde{\boldsymbol{w}}_k^{(i)}\}$ such that $\sum_{i=1}^{N}\tilde{\boldsymbol{w}}_k^{(i)}=1$
12: Sample N times with replacement a set of indices $\{i^-:i=1,\ldots,N\}$, from a distribution with $\mathbb{P}(i^-=j)=\tilde{\boldsymbol{w}}_k^{(j)}$
13: **for** $i=1,\ldots,N$ **do**
14: Sample a particle $^{(i)}_{k}\sim\mathcal{N}(\mathbf{m}_{k-1}^{(i^-)},h^{2-\eta}\mathbf{V}_{k-1})$, where $\mathbf{V}_{k-1}=\sum_{i=1}^{N}\boldsymbol{w}_{k-1}^{(i)}(\ ^{(i)}_{k-1}-\ ^{-}_{k-1})(\ ^{(i)}_{k-1}-\ ^{-}_{k-1})^T$
15: Assign weights $\boldsymbol{w}_k^{(i)}\propto\frac{p(z_k|\boldsymbol{\theta}_k^{(i)};\mathbf{x}_k)}{p(z_k|\mathbf{m}_{k-1}^{(i^-)};\mathbf{x}_k)}$
16: **end for**
17: Normalize $\{\boldsymbol{w}_k^{(i)}\}$ such that $\sum_{i=1}^{N}\boldsymbol{w}_k^{(i)}=1$
18: **end for**

4. ACTIVE SENSING

Algorithm 1 recursively computes approximations of the posterior pdfs, for given measurement locations $\mathbf{x}_k, k=0,1,\ldots,$ where in

[1] Note that methods similar to those in the area of *system identification with binary measurements* (see e.g. [16, 17]) can also be used to estimate $\beta_j, j=1,\ldots,J$. However, in general the performance is not as good as with SMC methods. Furthermore, it is not clear how to choose and optimize the measurement locations $\{\mathbf{x}_k\}$ using these methods.

Algorithm 2 Active sensing algorithm:
$\mathbf{x}_{k+1} = \texttt{ActiveSensing}(\mathbf{x}_k, \{\ ^{(i)}_k\})$

1: **Algorithm Parameters**: $\varepsilon \geq 0$, $\alpha \in [0,\infty)\backslash\{1\}$, $\rho_0 \geq 0$, $N_\rho \in \mathbb{N}$, $N_d \in \mathbb{N}$, search region $\mathcal{S}$
2: **Inputs**: $\mathbf{x}_k$, $\{\ ^{(i)}_k\}$
3: **Output**: Next measurement location $\mathbf{x}_{k+1}$
4: With prob. ε set $\mathbf{x}_{k+1}$ to a random location in $\mathcal{S}$, otherwise set

$$\mathbf{x}_{k+1} = \arg\max_{\mathbf{x}'\in\mathcal{X}_k} \frac{1}{\alpha-1}\sum_{z_{k+1}=0}^{1} \gamma_1(z_{k+1}|\mathbf{x}') \ln \frac{\gamma_\alpha(z_{k+1}|\mathbf{x}')}{(\gamma_1(z_{k+1}|\mathbf{x}'))^\alpha}$$

where $\mathcal{X}_k$ is given by (9) and γ_α is given by (8).

Algorithm 3 Parameter estimation with active sensing

1: **Algorithm Parameters**: $N \in \mathbb{N}$, $a \in (0,1)$, $h = \sqrt{1-a^2}$, $\eta \geq 0$, prior pdf $p_0(\ )$
2: **Inputs**: Initial location $\mathbf{x}_1$
3: **Outputs**: Particles $\{\ ^{(i)}_k\}$ and weights $\{\boldsymbol{w}_k^{(i)}\}$
4: Sample particles $^{(i)}_0, i = 1,\ldots,N$ from $p_0(\ )$, and assign weights $\boldsymbol{w}_0^{(i)} = \frac{1}{N}, i = 1,\ldots,N$
5: **for** $k = 1, 2, \ldots,$ **do**
6: Run lines 6 - 17 of Algorithm 1
7: Determine $\mathbf{x}_{k+1} = \texttt{ActiveSensing}(\mathbf{x}_k, \{\ ^{(i)}_k\})$ using Algorithm 2
8: **end for**

principle these locations can be arbitrary. In this section we want to adaptively choose the next measurement location $\mathbf{x}_{k+1}$ based on the measurements $\mathbf{x}_{1:k}$ collected so far, which we will refer to as *active sensing*, with the aim of achieving faster convergence of the parameter estimates. We will follow an approach to active sensing that is based on [7], see also [8, 13], while also incorporating a random exploration mode.

For a given posterior pdf $p(\ |z_{1:k};\mathbf{x}_{1:k})$, we want to select the next measurement location $\mathbf{x}_{k+1}$ to be the one which maximizes a one-step ahead expected reward function

$$\mathbf{x}_{k+1} = \arg\max_{\mathbf{x}'\in\mathcal{X}_k} \mathbb{E}[\mathcal{R}(\mathbf{x}', p(\ |z_{1:k};\mathbf{x}_{1:k}))], \tag{7}$$

where the reward function $\mathcal{R}((\mathbf{x}', p(\ |z_{1:k};\mathbf{x}_{1:k}))$ is chosen to provide a measure of the amount of additional information that can be obtained by choosing location $\mathbf{x}'$. Following [7,13], the reward function will be chosen as the Rényi divergence [20] between the posterior pdf $p(\ |z_{1:k};\mathbf{x}_{1:k})$ at iteration k and the future posterior pdf $p(\ |z_{1:k+1};\mathbf{x}_{1:k+1})$ at iteration $k+1$. The Rényi divergence between two pdfs $f_0(.)$ and $f_1(.)$ is defined as

$$D_\alpha(f_1||f_0) \triangleq \frac{1}{\alpha-1}\ln\int f_1^\alpha(\mathbf{t}) f_0^{1-\alpha}(\mathbf{t})d\mathbf{t},$$

where the parameter $\alpha \geq 0$ with $\alpha \neq 1$. When the posterior pdfs are approximated by particles $\{\ ^{(i)}_k\}$, with corresponding weights $\{\boldsymbol{w}_k^{(i)}\}$, it can be shown (see [7, 13]) that

$$\mathcal{R}(\mathbf{x}', p(\ |z_{1:k};\mathbf{x}_{1:k})) \approx \frac{1}{\alpha-1}\ln\frac{\gamma_\alpha(z_{k+1}|\mathbf{x}')}{(\gamma_1(z_{k+1}|\mathbf{x}'))^\alpha},$$

where

$$\gamma_\alpha(z_{k+1}|\mathbf{x}') \triangleq \frac{1}{N}\sum_{i=1}^{N} p(z_{k+1}|\ ^{(i)}_k;\mathbf{x}')^\alpha. \tag{8}$$

The expected reward can then be approximated by

$$\frac{1}{\alpha-1}\sum_{z_{k+1}=0}^{1}\gamma_1(z_{k+1}|\mathbf{x}')\ln\frac{\gamma_\alpha(z_{k+1}|\mathbf{x}')}{(\gamma_1(z_{k+1}|\mathbf{x}'))^\alpha}.$$

For problem (7), we also need to specify the set $\mathcal{X}_k$ of candidate future measurement locations. In general, $\mathcal{X}_k$ can depend on both $z_{1:k}$ and $\mathbf{x}_{1:k}$. Here we assume that an agent can/should only move a limited distance from its current position $\mathbf{x}_k$ in order to reach a new measurement location. Furthermore, for computational tractability in problem (7), we will restrict $\mathcal{X}_k$ to be a finite set. Recall that $\mathcal{S}$ is the allowable search area for the agent. Similar to [13], we let[2]

$$\mathcal{X}_k = \Big\{\mathbf{x}_k + \Big(n\rho_0\cos\Big(\frac{2\pi\ell}{N_d}\Big), n\rho_0\sin\Big(\frac{2\pi\ell}{N_d}\Big)\Big), \\ n = 0,\ldots,N_\rho, \ell = 0,1,\ldots,N_d-1\Big\}\cap\mathcal{S}, \tag{9}$$

where ρ_0 represents a radial step size, N_ρ the maximum number of step sizes between two consecutive measurement locations, and N_d the number of different directions the agent can move in the plane. The case $n = 0$ in (9) corresponds to the agent staying in its current position to collect another measurement. The cardinality of the set $\mathcal{X}_k$ thus satisfies $|\mathcal{X}_k| \leq N_\rho N_d + 1$.

Finally, we also introduce an exploration parameter ε [21], such that with probability $1-\varepsilon$ we carry out the above optimization (7), but with probability ε the next measurement location will simply be a random location in the search space $\mathcal{S}$. The reason for introducing this exploration parameter is that sometimes it is very difficult to tell whether there are any potential sources in unexplored areas, as the effects of sources on the field values usually decay with distance. The full parameter estimation scheme which includes active sensing is summarized in Algorithms 2 and 3.

5. SIMULATION STUDIES

5.1. Example 1

We consider first an example where the true field can be modelled using 4 basis functions, with parameters $\tilde{\beta}_1 = 1.3$, $\tilde{\beta}_2 = 0.4$, $\tilde{\beta}_3 = 1.0$, $\tilde{\beta}_4 = 0.7$, $\tilde{\mathbf{c}}_1 = (1,1)$, $\tilde{\mathbf{c}}_2 = (1,2)$, $\tilde{\mathbf{c}}_3 = (2,2)$, $\tilde{\mathbf{c}}_4 = (2,1)$, and $\tilde{\sigma}_1^2 = \tilde{\sigma}_2^2 = \tilde{\sigma}_3^2 = \tilde{\sigma}_4^2 = 0.5$, where the tilde is used to denote the true parameters. The search region is taken to be $\mathcal{S} = [0,3]\times[0,3]$. The measurement noise variance $\sigma_v^2 = 0.1$, and the sensor threshold $\tau = 1$.

In this first example we assume that the true values of $\tilde{\mathbf{c}}_j, \tilde{\sigma}_j^2, j = 1,\ldots,4$ are known and set $\mathbf{c}_j, \sigma_j^2$ to the above values, while $\beta_j, j = 1,\ldots,4$ and σ_v^2 are to be estimated. We will compare the performance without (Algorithm 1) and with (Algorithm 3) active sensing. For Algorithm 1, we use $N = 5000$ particles, $a = 0.8$, $\eta = 0.01$. The prior pdf $p_0(\ )$ is chosen where components of corresponding to β_j are uniformly distributed between 0 and 2, and the component corresponding to $\log\sigma_v$ is uniformly

[2] One can also incorporate other features of the agent vehicle dynamics by appropriately modifying $\mathcal{X}_k$, e.g. by restricting the set of turn angles of the agent, though for simplicity of presentation we will not do this here.

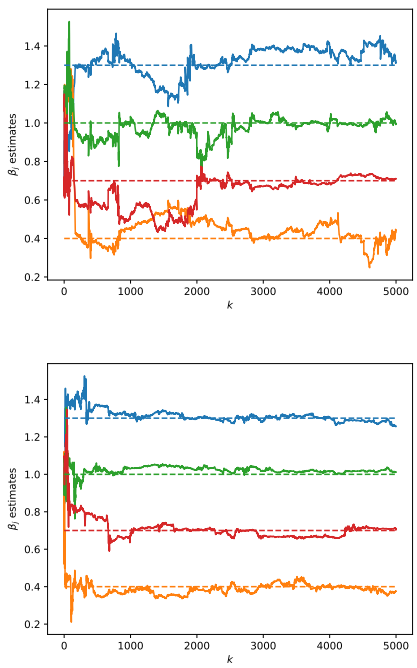


Fig. 1: Example 1 - Estimates of $\beta_1, \ldots, \beta_4$ using Algorithm 1 (top) and Algorithm 3 (bottom)

distributed between $\log 0.01$ and $\log 0.5$. The measurement locations at every iteration are chosen in a similar manner to a 2-dimensional random walk as follows: The initial position $\mathbf{x}_1$ is set at $(1.5, 1.5)$, and $\mathbf{x}_{k+1} = \mathbf{x}_k + (\rho_k \cos\psi_k, \rho_k \sin\psi_k)$ if $\mathbf{x}_k + (\rho_k \cos\psi_k, \rho_k \sin\psi_k) \in \mathcal{S}$, where ρ_k is uniformly distributed between 0 and 0.2, and ψ_k is uniformly distributed between 0 and 2π. If $\mathbf{x}_k + (\rho_k \cos\psi_k, \rho_k \sin\psi_k) \notin \mathcal{S}$, then we stay in the current position, i.e. $\mathbf{x}_{k+1} = \mathbf{x}_k$. For Algorithm 3, we use the same N, a, η, and prior pdf $p_0(\ )$. Additionally, we use exploration parameter $\varepsilon = 1/100$, Rényi divergence parameter $\alpha = 1/2$, and $\rho_0 = 0.1$, $N_\rho = 1$, $N_d = 8$ for the search space $\mathcal{X}_k$. The initial position is $\mathbf{x}_1 = (1.5, 1.5)$. On a Core i7-9700 desktop with 16 Gb of RAM, each iteration of Algorithm 3 takes less than 10 ms.

In Fig. 1 we plot the conditional mean estimates of $\beta_1, \ldots, \beta_4$ over different iterations. We see that the parameter estimates using the active sensing approach will become close to and stabilize around the true values in significantly fewer iterations than for Algorithm 1. The variations in the estimates also seems to be reduced when compared to Algorithm 1 (for the same values of a and η).

5.2. Example 2

In this example, the true field is represented using 4 basis functions, with randomly generated parameters $\tilde{\beta} = 1.39$, $\tilde{\beta}_2 = 1.14$, $\tilde{\beta}_3 = 0.85$, $\tilde{\beta}_4 = 1.08$ $\tilde{\mathbf{c}}_1 = (0.45, 2.18)$, $\tilde{\mathbf{c}}_2 = (0.84, 0.56)$, $\tilde{\mathbf{c}}_3 = (1.06, 0.42)$, $\tilde{\mathbf{c}}_4 = (1.95, 2.30)$, $\tilde{\sigma}_1^2 = 1.06$, $\tilde{\sigma}_2^2 = 0.71$, $\tilde{\sigma}_3^2 = 0.59$, $\tilde{\sigma}_4^2 = 1.00$, where the tilde is again used to denote the true parameters. A plot of the true field is shown in Fig. 2. The search region is $\mathcal{S} = [0, 3] \times [0, 3]$, the measurement noise variance is $\sigma_v^2 = 0.1$, and the sensor threshold is $\tau = 1$.

The true values of $\tilde{c}_j, \tilde{\sigma}_j^2, j = 1, \ldots, 4$ are not known in this example. We consider using $J = 4$ and $J = 16$ basis functions, with

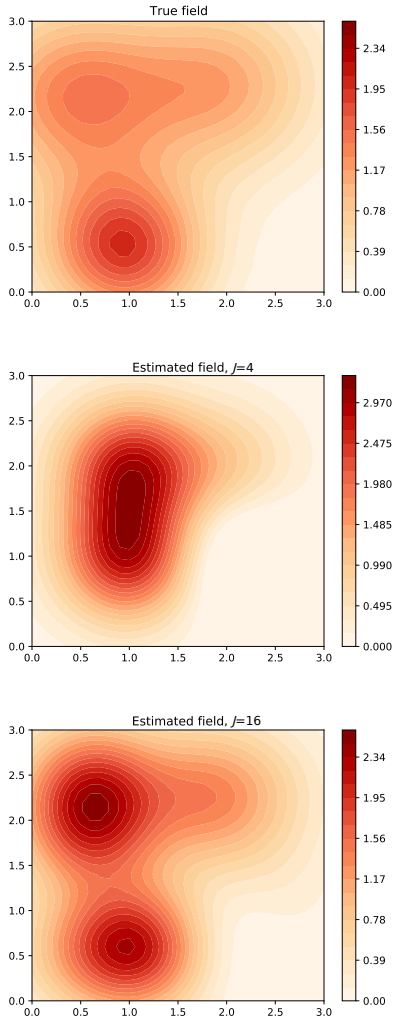


Fig. 2: Example 2 - True field (top), and estimated fields using 4 (middle) and 16 (bottom) basis functions

$\mathbf{c}_j$'s equally spaced in $[0, 3] \times [0, 3]$ (where the x and y coordinates are located at $\frac{3m}{\sqrt{J}+1}, m = 1, \ldots, \sqrt{J}$) and $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 0.5$. Plots of the estimated fields using parameter estimates obtained after 5000 iterations of Algorithm 3 are shown in Fig. 2. We see that for $J = 4$, the reconstructed field qualitatively is not particularly close to the true field when we don't know the true $\mathbf{c}_j$'s and σ_j^2's. However, as the number of basis functions used is increased, the field can be more accurately approximated.

6. CONCLUSION

This paper has considered the problem of field estimation using an autonomous mobile agent with binary sensor measurements. Algorithms based on sequential Monte Carlo techniques have been presented. The incorporation of active sensing for choosing the measurement locations has been demonstrated to reduce the time required to estimate the field. Extensions of this work to time-varying fields and multiple agents are currently under investigation.

7. REFERENCES

[1] A. J. Lilienthal, M. Reggente, M. Trincavelli, J. L. Blanco, and J. Gonzalez, "A statistical approach to gas distribution modelling with mobile robots - The kernel DM+V algorithm," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, St. Louis, USA, Oct. 2009, pp. 570–576.

[2] A. Gunatilaka, B. Ristic, and R. Gailis, "On localisation of a radiological point source," in *Proc. Information, Decision and Control*, Adelaide, Australia, Feb. 2007, pp. 236–241.

[3] H. M. La, W. Sheng, and J. Chen, "Cooperative and active sensing in mobile sensor networks for scalar field mapping," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 1, pp. 1–12, Jan. 2015.

[4] M. R. Morelande and A. Skvortsov, "Radiation field estimation using a Gaussian mixture," in *Proc. Intl. Conf. Inf. Fusion*, Seattle, USA, July 2009, pp. 2247–2254.

[5] R. A. Razak, S. Sukumar, and H. Chung, "Estimating scalar fields with mobile sensor networks," in *Proc. Indian Control Conf.*, Hyderabad, India, Dec. 2019, pp. 63–66.

[6] A. Doucet, N. de Freitas, and N. Gordon, "Sequential Monte Carlo methods in practice," Springer, New York, 2001.

[7] C. M. Kreucher, A. O. Hero, K. D. Kastella, and M. D. Morelande, "An information-based approach to sensor management in large dynamic networks," *Proc. IEEE*, vol. 95, no. 5, pp. 978–999, May 2007.

[8] M. Morelande and B. Ristic, "Detection and estimation of radiological sources," in *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*, M. Mallick, V. Krishnamurthy, and B.-N. Vo, Eds., chapter 15, pp. 579–616. John Wiley & Sons, 2013.

[9] K. Rosser, K. Pavey, N. FitzGerald, A. Fatiaki, D. Neumann, D. Carr, B. Hanlon, and J. Chahl, "Autonomous chemical vapour detection by micro UAV," *Remote Sensing*, vol. 7, pp. 16865–16882, 2015.

[10] P. P. Neumann, V. Hernandez Bennetts, A. J. Lilienthal, M. Bartholmai, and J. H. Schiller, "Gas source localization with a micro-drone using bio-inspired and particle filter-based algorithms," *Advanced Robotics*, vol. 27, no. 9, pp. 725–738, 2013.

[11] G. Battistelli, L. Chisci, N. Forti, and S. Gherardini, "MAP moving horizon estimation for threshold measurements with application to field monitoring," *Int. Journal Adapt. Control and Signal Process.*, pp. 1–16, 2019.

[12] D. D. Selvaratnam, I. Shames, D. V. Dimarogonas, J. H. Manton, and B. Ristic, "Co-operative estimation for source localisation using binary sensors," in *Proc. IEEE Conf. Decision and Control*, Melbourne, Australia, Dec. 2017, pp. 1572–1577.

[13] B. Ristic, M. Morelande, and A. Gunatilaka, "Information driven search for point sources of gamma radiation," *Signal Processing*, vol. 90, pp. 1225–1239, 2010.

[14] J. Park and I. W. Sandberg, "Approximation and radial-basis-function networks," *Neural Computation*, vol. 5, no. 2, pp. 305–316, 1993.

[15] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds., chapter 10, pp. 197–223. Springer, New York, 2001.

[16] L. Y. Wang, G. G. Yin, J.-F. Zhang, and Y. Zhao, *System Identification with Quantized Observations*, Birkhauser, Boston, 2010.

[17] K. Jafari, J. Juillard, and E. Colinet, "A recursive system identification method based on binary measurements," in *Proc. IEEE Conf. Decision and Control*, Atlanta, USA, Dec. 2010, pp. 1157–1158.

[18] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, New York, 2nd edition, 2004.

[19] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[20] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Math. Statist. and Prob.*, Berkeley, USA, 1961, pp. 547–567.

[21] R. S. Sutton and A. G. Barto, *Reinforcement Learning*, The MIT Press, Massachusetts, 2nd edition, 2018.

MULTICHANNEL NONNEGATIVE MATRIX FACTORIZATION WITH MOTOR DATA-REGULARIZED ACTIVATIONS FOR ROBUST EGO-NOISE SUPPRESSION

Alexander Schmidt and Walter Kellermann

Multimedia Communications and Signal Processing,
Friedrich-Alexander University Erlangen-Nürnberg,
Cauerstr. 7, 91058 Erlangen, Germany,
{alexander.as.schmidt, walter.kellermann}@fau.de

ABSTRACT

The suppression of ego-noise is often addressed using dictionary-based methods where the characteristic spectral structure of ego-noise is approximated by a linear combination of dictionary entries. A blind, entirely audio data-based selection of the dictionary entries is, however, challenging and reacts sensitive against other signals besides ego-noise in a mixture. For a more robust behavior, we propose a motor data-dependent regularization term which promotes similar activations for similar physical states of the robot. The proposed regularization term is added to a multichannel nonnegative matrix factorization (MNMF)-based signal model and according update rules are derived. We analyze the proposed method for a challenging ego-noise scenario and demonstrate the efficacy of the method compared to an approach for which no motor data is used.

***Index Terms*—** Ego-noise, multichannel nonnegative matrix factorization, motor data

1. INTRODUCTION

The acoustic domain carries a wealth of information about the surrounding scene of an autonomous system (AS). An efficient and robust acoustic scene analysis (ASA) can therefore significantly improve the perception of an AS and help augmenting self-awareness (cf. acoustic self-awareness [1]). As a major challenge for robust ASA, a microphone-equipped AS is typically exposed to a variety of different noise sources, especially and very specific for ASs, ego-noise, i.e., self-created noise, which is is caused by the mechanical and electric components of an AS. In this contribution, we consider ego-noise of a humanoïd robot as special kind of an AS. Since the joints of the robot are located near the microphones, the Signal-to-Noise ratio (SNR) between a signal of interest and ego-noise is typically low. Besides, humanoïd's ego-noise is highly non-stationary due to the highly irregular movements of the robot.

For suppression approaches, it is usually exploited that ego-noise exhibits a characteristic structure in the short-time Fourier transform (STFT) domain which cannot be arbitrarily complex due to the limited number of degrees of freedom of the robot. This fact motivates the use of learning-based dictionary methods for single-channel ego-noise suppression as nonnegative matrix factorization (NMF) [2, 3, 4]. Recently, multichannel dictionary approaches have been proposed such as phase-optimized PO-KSVD [5]) and multichannel nonnegative matrix factorization (MNMF) [6]. The latter method will be extended in this work.

Additionally, ego-noise suppression can benefit from non-acoustic reference information (NARI) describing the internal physical state of the robot. A promising kind of NARI is motor state information which will be referred to as motor data in the following. In several approaches, motor data is used to estimate ego-noise power spectral densities (PSDs) which are subsequently used for spectral subtraction. In [7, 8], this is achieved using a data set of noise templates, while in [9] an artificial neural network is considered for estimation. In [10], it was shown that the harmonic structure of ego-noise can be estimated using motor data. This knowledge can be incorporated into the dictionary which was demonstrated to be beneficial for ego-noise modeling.

In [11], the update rules of single-channel NMF were extended by a motor data-dependent regularization term which promotes similar activations of dictionary entries in time frames in which similar motor data samples are measured. By this, a more robust estimation of activations was achieved. In this contribution, this approach is extended to multiple microphones employing an MNMF signal model. We first derive a graph structure which encodes the similarity between motor data samples. From this, a regularization term is derived which is incorporated into the MNMF cost function and according update rules are derived.

This paper is structured as follows. In Sec. 2.1, we introduce a signal model which is subsequently used to describe the fundamental concept of MNMF in Sec. 2.2. In Sec. 2.3, the motor data-dependent regularization term is derived and added to the MNMF cost function. Experimental results in Sec. 3 demonstrate the efficacy of the derived update rules.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

2. MOTOR DATA-REGULARIZED MNMF

In this section, we first introduce an appropriate signal model, review conventional MNMF and finally present the proposed motor data-dependent regularization term.

2.1. Signal model

We consider a humanoïd robot which is equipped with I microphones. We denote the STFT representation of the i-th microphone channel as $\mathbf{Y}^{(i)} \in \mathbb{C}^{N\times M}$, where N is the number of frequency bins and M is the number of time frames. We concatenate all I channels for a single time-frequency bin nm and introduce $\mathbf{Y}_{nm} = [Y_{nm}^{(1)}, \ldots, Y_{nm}^{(I)}]^{\mathrm{T}} \in \mathbb{C}^I$, where $Y_{nm}^{(i)}$ denotes the nm-th bin of $\mathbf{Y}^{(i)}$, $i = 1, \ldots, I$, and $\cdot^{\mathrm{T}}$ is the transpose operator. Based on this, we define the spatial correlation matrix as $\mathbf{\Phi}_{nm} = \mathcal{E}\left\{\mathbf{Y}_{nm}\mathbf{Y}_{nm}^{\mathrm{H}}\right\}$, where $\mathcal{E}\{\cdot\}$ denotes the expectation operator and $\cdot^{\mathrm{H}}$ the Hermitian operator. We assume that for every time frame m a motor data vector $\boldsymbol{\alpha}_m$ is available which characterizes the physical state of the robot in this time frame. $\boldsymbol{\alpha}_m$ contains the angular position data collected by the proprioceptors of the robot's joints and, additionally, the instantaneously estimated angular velocity per joint. Thus, if the robot is equipped with L proprioceptors, we have $\boldsymbol{\alpha}_m \in \mathbb{R}^{2L}$. A detailed description of the motor data measurement and construction of $\boldsymbol{\alpha}_m$ is provided in [11].

2.2. Multichannel NMF (MNMF)

First, we review briefly the fundamental idea of single-channel NMF and subsequently introduce MNMF as extension. For NMF, the element-wise squared magnitude of $\mathbf{Y}^{(i)}$, $|\mathbf{Y}^{(i)}|^2$, is approximated by the matrix product $|\mathbf{Y}^{(i)}|^2 \approx \mathbf{D}^{(i)}\mathbf{H}^{(i)}$, where $\mathbf{D}^{(i)} \in \mathbb{R}_+^{N\times K}$ and $\mathbf{H}^{(i)} \in \mathbb{R}_+^{K\times M}$ are referred to as dictionary and activation matrix, respectively. K is the dictionary size. To derive the Itakura-Saito (IS) divergence-based NMF cost function, it is assumed that all elements of $\mathbf{Y}^{(i)}$ are statistically independent and distributed according to a complex Gaussian distribution with variance $\phi_{nm}^{(i)} = |\mathbf{Y}_{nm}^{(i)}|^2$. Minimizing the log-likelihood with respect to the unknown model parameters results in iterative update rules for $\mathbf{D}^{(i)}, \mathbf{H}^{(i)}$ [12]. The Maximum Likelihood (ML) interpretation of NMF motivates its multichannel extension. It is assumed that the multichannel observations $\mathbf{Y}_{nm}$ are distributed according to a multivariate complex Gaussian

$$\mathcal{N}_{\mathrm{c}}(\mathbf{Y}_{nm}|0, \mathbf{\Phi}_{nm}) \sim \frac{1}{\det \mathbf{\Phi}_{nm}} \exp\left(-\mathbf{Y}_{nm}^{\mathrm{H}}\mathbf{\Phi}_{nm}^{-1}\mathbf{Y}_{nm}\right),$$

where $\mathbf{\Phi}_{nm}$ is the spatial correlation matrix of $\mathbf{Y}_{nm}$. Again assuming statistical independence of all time-frequency bins, the log-likelihood can be written as

$$\begin{aligned}\log\ p(\mathbf{Y}|\mathbf{\Phi}) &= \log \prod_{n=1}^{N}\prod_{m=1}^{M} \mathcal{N}_{\mathrm{c}}(\mathbf{Y}_{nm}|0, \mathbf{\Phi}_{nm}) \\ &= \sum_{n=1}^{N}\sum_{m=1}^{M}\left(-\log\det\mathbf{\Phi}_{nm} - \mathrm{tr}(\tilde{\mathbf{\Phi}}_{nm}\mathbf{\Phi}_{nm}^{-1})\right)\end{aligned} \tag{1}$$

with $\tilde{\mathbf{\Phi}}_{nm} = \mathbf{Y}_{nm}\mathbf{Y}_{nm}^{\mathrm{H}}$ and trace operator $\mathrm{tr}(\cdot)$. Note that $\mathbf{Y} \in \mathbb{C}^{I\times N\times M}$ and $\mathbf{\Phi} \in \mathbb{C}^{I\times I\times N\times M}$ denote tensors. By negating Eq. 1, we can derive the MNMF cost function

$$J_{\mathrm{MNMF}} = \sum_{n=1}^{N}\sum_{m=1}^{M}\left(\mathrm{tr}(\tilde{\mathbf{\Phi}}_{nm}\mathbf{\Phi}_{nm}^{-1}) + \log\det\mathbf{\Phi}_{nm}\right). \tag{2}$$

The spatial correlation matrix $\mathbf{\Phi}_{nm}$ is modeled as [13]

$$\hat{\mathbf{\Phi}}_{nm} = \sum_{k=1}^{K}\mathbf{G}_{nk}d_{nk}h_{km}, \tag{3}$$

where $\mathbf{G}_{nm}$ is a Hermitian spatial transfer matrix and d_{nk} and h_{km} are the nk-th and km-th entries of $\mathbf{D}$ and $\mathbf{H}$, respectively. For the considered application of ego-noise suppression, Eq. 3 can be interpreted as K ego-noise sources, where source k has a distinct spectral structure d_{nk} and is time-dependently activated by h_{km}. Additionally, every source k possesses a different spatial characteristic $\mathbf{G}_{nk}$. It was shown in [6] that Eq. 3 indeed represents an appropriate model for ego-noise which is caused by various mechanical components distributed over the robot's body. The cost function defined by Eq. 2 and Eq. 3 can be minimized by updating the unknown model parameters $\mathbf{D}$, $\mathbf{H}$ and $\{\mathbf{G}_{nk}\}_{\forall n,k}$ iteratively. According multiplicative update rules can be derived employing the Majorize-Minimization (MM) algorithm [13].

In this paper, we employ ego-noise suppression in a semi-supervised manner (cf. Sec. 2.4). Using recordings which contain ego-noise only, we learn a ego-noise dictionary and according spatial correlation matrices. This model is subsequently used to model the ego-noise within a mixture by estimating activations while keeping the dictionary and spatial correlation matrices fixed. The estimation of the activations is sensitive regarding other signals in the mixture, e.g., speech. To increase robustness, we propose to add a motor data-dependent regularization term to Eq. 2.

2.3. Motor data-driven regularization term

The fundamental idea of the proposed regularization term is to enforce similar activations for time instances in which the state of the robot, and hence the emitted ego-noise, is similar. To measure the similarity of robot states in frames m and m', we compare motor data vectors $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_{m'}$ by evaluating a Gaussian kernel

$$W_{mm'} = \exp\left(-\|\boldsymbol{\alpha}_m - \boldsymbol{\alpha}_{m'}\|_2^2 \,/\, 2\epsilon^2\right) \in (0, 1], \tag{4}$$

with squared Euclidean norm $\|\cdot\|_2^2$ and scale parameter $\epsilon \in \mathbb{R}_+$. The larger $W_{mm'}$, the larger the affinity between $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_{m'}$ is. Note that $W_{mm'} = 1$ if $\boldsymbol{\alpha}_m = \boldsymbol{\alpha}_{m'}$. The pairwise computation of the affinity according to Eq. 4 can be interpreted as constructing an undirected graph where $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M$ represent the nodes and $W_{mm'}$ is the weight between nodes $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_{m'}$. The connectivity of the graph can be controlled with ϵ. The regularization term is then

constructed as

$$\mathcal{R} = \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{m'=1}^{M} \mathrm{d}\left(h_{km} | h_{km'}\right) \cdot W_{mm'} \tag{6}$$

where $\mathrm{d}(\cdot|\cdot)$ is a distance function which is symmetric with respect to (w.r.t.) its input arguments and is minimal for equal activations $h_{km} = h_{km'}$. If two motor data vectors $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_{m'}$ are similar, $W_{mm'}$ in Eq. 6 is close to one. Hence, minimizing Eq. 6 w.r.t. h_{km} enforces similarity between h_{km} and $h_{km'}$. A natural choice for $\mathrm{d}(\cdot|\cdot)$ is, e.g., the Euclidean distance. However, our experiments showed that an improved ego-noise suppression is achieved using

$$\begin{aligned} \mathrm{d}\left(h_{km} | h_{km'}\right) &= \frac{1}{2}(\mathrm{d}_{\mathrm{IS}}\left(h_{km} | h_{km'}\right) + \mathrm{d}_{\mathrm{IS}}\left(h_{km'} | h_{km}\right)) \\ &= \frac{1}{2}\left(\frac{h_{km}}{h_{km'}} + \frac{h_{km'}}{h_{km}} - 2\right), \end{aligned} \tag{7}$$

where $\mathrm{d}_{\mathrm{IS}}\left(a|b\right) = a/b - \log(a/b) - 1,\ a, b \in \mathbb{R}_+$ is the IS divergence. While the IS divergence is non-symmetric, the sum of $d_{\mathrm{IS}}(a|b)$ and $d_{\mathrm{IS}}(b|a)$ is symmetric again. The proposed novel regularized MNMF (R-MNMF) cost function reads

$$J_{\text{R-MNMF}} = J_{\text{MNMF}} + \lambda \cdot \mathcal{R}, \tag{8}$$

where $\lambda \geq 0$ controls the influence of the regularization term. As outlined in Sec. 2.2, the iterative update rules minimizing Eq. 2 are derived using the MM algorithm: first, a majorizer for Eq. 2 is constructed which is of simpler mathematical structure such that it can be more easily minimized subsequently. The majorizer for Eq. 2 can be found in [13]. The update terms for the novel cost function Eq. 8 are derived similarly as follows: We first note that Eq. 6 does not depend on d_{nk} or $\mathbf{G}_{nk}$. Therefore, only the update rule for h_{km} is affected by the motor data-dependent regularization term. Second, since the regularization term is nonnegative, we can straightforwardly obtain an upper bound for Eq. 8 if we add the derived regularization term to the upper bound of the original cost function Eq. 2. The obtained term is minimized w.r.t. to h_{kn}. The result is given in Eq. 5, where $\leftarrow$ illustrates that Eq. 5 is an update equation, i.e., the novel estimate for h_{kn} depends on h_{kn} obtained in the previous iteration. For $\lambda = 0$, the conventional MNMF update term for h_{kn} is obtained.

2.4. Ego-noise suppression

We employ a semi-supervised approach for ego-noise suppression [15]. Given audio data containing ego-noise only and associated motor data, we train an ego-noise model consisting of an ego-noise dictionary $\mathbf{D}_{\text{EN}}$, activation matrix $\mathbf{H}_{\text{EN}}$ and a set of spatial correlation matrices $\{\mathbf{G}_{\text{EN},nk}\}_{\forall n,k}$. Note that the required ego-noise for training could be recorded, e.g., during an initial configuration phase of the robot simultaneous to a calibration of other sensors. In the following suppression step, we consider a mixture containing ego-noise and a desired speech signal. Additionally, also motor data is available again. We initialize another dictionary $\mathbf{D}_{\text{S}}$ with activation matrix $\mathbf{H}_{\text{S}}$ and associated spatial correlation matrices $\{\mathbf{G}_{\text{S},nk}\}_{\forall n,k}$ to model the speech signal. Then, $\mathbf{D}_{\text{S}}$, $\mathbf{H}_{\text{S}}$, $\{\mathbf{G}_{\text{S},nk}\}_{\forall n,k}$ and $\mathbf{H}_{\text{EN}}$ are updated while $\mathbf{D}_{\text{EN}}$ and $\{\mathbf{G}_{\text{EN},nk}\}_{\forall n,k}$ are kept constant. We obtain a speech estimate by applying a multichannel Wiener filter (MWF) [16, 17]

$$\hat{\mathbf{Y}}_{\text{S},nm} = \hat{\boldsymbol{\Phi}}_{\text{S},nm} \left(\hat{\boldsymbol{\Phi}}_{\text{S},nm} + \hat{\boldsymbol{\Phi}}_{\text{EN},nm}\right)^{-1} \mathbf{Y}_{nm} \tag{9}$$

where $\hat{\boldsymbol{\Phi}}_{\text{S},nm}$ and $\hat{\boldsymbol{\Phi}}_{\text{EN},nm}$ can be computed using Eq. 3 using $\mathbf{D}_{\text{S}}$, $\mathbf{H}_{\text{S}}$, $\{\mathbf{G}_{\text{S},nk}\}_{\forall n,k}$ and $\mathbf{D}_{\text{EN}}$, $\mathbf{H}_{\text{EN}}$, $\{\mathbf{G}_{\text{EN},nk}\}_{\forall n,k}$.

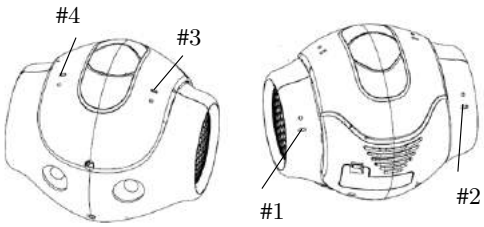


Fig. 1. Microphone positions of the humanoïd robot NAO (image adapted from [14]).

3. RESULTS

For evaluation, we conducted experiments with a NAO™ humanoïd robot [18] which is equipped with four microphones mounted to its head, cf. Fig. 1. For all experiments, all four microphones were used, i.e., $I = 4$. The robot has 26 joints: 2 in the head, 12 in the arms and 12 in the legs. Due to the immediate proximity to the microphones, the arm joints generate ego-noise which is much louder than the noise originating from movements of the legs. In the following, we concentrate on the more challenging scenario and consider therefore exclusively ego-noise from arm movements. Specifically, the right and left arm are moved with randomly chosen trajectories and speeds. The movements were initiated randomly which results in highly diverse ego-noise originating from ei-

$$h_{km} \leftarrow \sqrt{\frac{\sum_{n=1}^{N} \mathrm{tr}\left(\hat{\boldsymbol{\Phi}}_{nm}^{-1} \tilde{\boldsymbol{\Phi}}_{nm} \hat{\boldsymbol{\Phi}}_{nm}^{-1} \mathbf{H}_{nk}\right) d_{nk} h_{km}^2 + \frac{\lambda}{2} \cdot \sum_{\substack{m'=1 \\ m' \neq m}}^{M} h_{km'} W_{mm'}}{\sum_{n=1}^{N} \mathrm{tr}\left(\tilde{\boldsymbol{\Phi}}_{nm}^{-1} \mathbf{H}_{nk}\right) d_{nk} + \frac{\lambda}{2} \cdot \sum_{\substack{m'=1 \\ m' \neq m}}^{M} \frac{1}{h_{km'}} W_{mm'}}} \tag{5}$$

Table 1. SIR, SDR (both in dB) and PESQ achieved by the proposed method, audio only-based MNMF and NMF for different SNRs of the input mixture. For PESQ, we provide a mean value averaged over all channels ($\overline{\text{PESQ}}$) and the peak value (pPESQ) where the channel achieving this score is given in brackets.

	SNR = -3dB				SNR = 0dB				SNR = 3dB			
	Unproc.	NMF	MNMF	R-MNMF $\lambda = 0.2$	Unproc.	NMF	MNMF	R-MNMF $\lambda = 0.4$	Unproc.	NMF	MNMF	R-MNMF $\lambda = 0.8$
SIR	-8.4	4.3	10.0	**12.7**	-5.6	7.5	13.6	**15.5**	-2.7	10.3	16.9	**17.7**
SDR	-8.8	1.9	6.8	**8.3**	-5.8	3.2	**9.6**	9.5	-2.8	4.2	**11.9**	9.3
$\overline{\text{PESQ}}$	1.1	1.2	1.5	**1.7**	1.1	1.3	1.82	**2.1**	1.2	1.3	2.2	**2.6**
pPESQ	1.1 (4)	1.2 (4)	1.9 (4)	**2.2** (4)	1.2 (2)	1.4 (4)	2.4 (4)	**2.6** (4)	1.3 (2)	1.5 (4)	2.9 (4)	**3.2** (4)

ther one of the two arms or both. Overall, all 12 joints of both arms were used, i.e., the resulting motor data vectors have dimension 24, i.e., $\boldsymbol{\alpha}_m \in \mathbb{R}^{24}$. We recorded 90s of ego-noise from which 40s were used for the training of a noise model. Experiments have shown that this amount of training data is sufficient to obtain overall satisfactory suppression results. The remaining 50s were employed for evaluating the algorithm. For this, we separately recorded speech utterances from the GRID corpus [19] (200 utterances in total, 18 male and 16 female speakers). The loudspeaker was positioned at 1m distance and 1.5m height of the robot's microphones. During all recordings, the cooling fan of the robot was switched off. Due to its highly stationary nature, the fan noise could, however, be easily suppressed using a MWF [20]. We added the recorded speech utterances to ego-noise which was not contained in the training. By varying the power of the desired speech signal, we created mixtures with SNR $\in \{-3, 0, 3\}$ dB measured in the front left microphone of the robot (channel 4, cf. Fig. 1). All recordings were conducted in a room with moderate reverberation ($T_{60} = 200$ms) and the height of the robot's microphones was kept constant at 55cm. The audio signals were sampled at $f_{\text{S}} = 16$kHz and transformed to the STFT domain using a Hamming window of length 64ms and an overlap of 50%.

As evaluation criteria, we use Signal-to-Distortion ratio (SDR) and Signal-to-Inference ratio (SIR) using Matlab functions provided by [21]. Note that both measures are averaged over the microphone channels. While SIR measures the overall suppression of ego-noise, SDR takes also the distortion of the desired speech signal by the suppression algorithm into account. To assess the subjective noise suppression performance, we additionally evaluate PESQ (perceptual evaluation of speech quality [22]) which generally reacts sensitively to artifacts introduced to the enhanced signal. We provide a mean PESQ value averaged over all channels and the highest PESQ of all channels. To demonstrate the benefit of the proposed method, we compare R-MNMF with conventional MNMF which uses audio signals only (referred to as audio-only MNMF in the following). Additionally, we provide the averaged suppression performance for single-channel NMF which is applied separately to each of the four microphone channels.

For R-MNMF and MNMF, the size of the ego-noise and the speech dictionary is chosen to $K_{\text{EN}} = 30$ and $K_{\text{S}} = 10$, respectively. For NMF, $K_{\text{EN}} = 30$ and $K_{\text{S}} = 20$ is set. All dictionary sizes were set such that they achieve satisfactory suppression results for all mixtures. For, R-MNMF, we chose $\epsilon = 0.02$ for all experiments. For training, the regularization parameter was set to $\lambda = 0.9$. During evaluation, λ is chosen SNR-dependent. It is given for every experiment separately. The chosen parameters have shown best suppression performance in terms of PESQ for the respective method. Results are provided in Table 1.

Typically, NMF introduces artifacts to the processed signal which is reflected in the low PESQ. Comparing the proposed method and audio-only MNMF, it can be observed that R-MNMF outperforms MNMF in terms of SIR for all SNRs. Regarding SDR, audio-only MNMF shows slight superior performance for SNRs 0dB and 3dB. This, however, is not reflected in the perceived quality of the enhanced signal. Both, average and peak PESQ show consistently superior behavior for proposed R-MNMF. Due to its position at the front of the robot's upper head (cf. Fig. 1), the signal of the fourth microphone channel is less affected by interfering ego-noise. Consequently, the highest peak PESQ was almost always achieved in the fourth microphone channel. Note that by choosing $\lambda = 0.9$ during the training phase of R-MNMF, the motor data-based regularization term implicitly influences the structure of the dictionary. During testing, λ must be increased for larger SNR to obtain the best result. This can be explained by the fact that the additional speech signal increasingly impairs the estimation of the activations for growing SNRs such that the influence of the regularization term must be chosen larger.

4. SUMMARY

In this paper, we introduced a novel approach for motor data-driven multichannel ego-noise suppression. We extended the cost function of MNMF with a motor data-dependent regularization term which enforces similar activations for time steps in which the robot is in similar physical states. We derived update terms and demonstrated the superiority of the proposed method for ego-noise from complex arm movements.

5. REFERENCES

[1] A. Schmidt, H. W. Löllmann, and W. Kellermann, "Acoustic self-awareness of autonomous systems in a world of sounds," *Proc. IEEE*, vol. 108, no. 7, pp. 1127–1149, 2020.

[2] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Proc. Advances Neural Inform. Process. Syst. (NIPS)*, vol. 13, 2001.

[3] C. Févotte and J. Idier, "Algorithms for non-negative matrix factorization with the β-divergence," *Neural Comput.*, vol. 23, no. 9, pp. 2421–2456, 2011.

[4] T. Tezuka, T. Yoshida, and K. Nakadai, "Ego-motion noise suppression for robots based on Semi-Blind Infinite Non-negative Matrix Factorization," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, Florence, Itlay, May 2014, pp. 6293–6298, IEEE.

[5] A. Deleforge and W. Kellermann, "Phase-optimized K-SVD for signal extraction from underdetermined multichannel sparse mixtures," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, Brisbane, QLD, Australia, Apr. 2015, pp. 355–359, IEEE.

[6] Thomas Haubner, A. Schmidt, and W. Kellermann, "Multichannel Nonnegative Matrix Factorization for Ego-Noise Suppression," in *Proc. ITG Fachtagung Sprachkommunikation*, Oldenburg, Germany, Oct. 2008, pp. 136–140, VDE-Verlag.

[7] G. Ince, K. Nakadai, T. Rodemann, Y. Hasegawa, H. Tsujino, and J. Imura, "Ego-noise suppression of a robot using template subtraction," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, St. Louis, MO, Oct. 2009, pp. 199–204, IEEE.

[8] G. Ince, K. Nakamura, F. Asano, H. Nakajima, and K. Nakadai, "Assessment of general applicability of ego noise estimation," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, San Francisco, CA, May 2011, pp. 3517–3522, IEEE.

[9] A. Ito, T. Kanayama, M. Suzuki, and S. Makino, "Internal noise suppression for speech recognition by small robots," in *Proc. European Conf. Speech Commun. and Technology (INTERSPEECH - Eurospeech)*, Lisbon, Portugal, 2005, pp. 2685–2688, Int. Speech Communication Assoc.

[10] A. Schmidt and W. Kellermann, "Informed Ego-noise Suppression Using Motor Data-driven Dictionaries," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, Brighton, UK, May 2019, pp. 116–120, IEEE.

[11] A. Schmidt, A. Brendel, T. Haubner, and W. Kellermann, "Motor data-regularized NMF for ego-noise suppression," *EURASIP J. Audio, Speech, Music Proc.*, vol. 11, no. 4, 2020.

[12] C. Févotte, N. Bertin, and J.-L. Durrieu, "Nonnegative Matrix Factorization with the Itakura-Saito Divergence: With Application to Music Analysis," *Neural Computation*, vol. 21, no. 3, pp. 793–830, Mar. 2009.

[13] H. Sawada, H. Kameoka, A. Araki, and N. Ueda, "Multichannel Extensions of NMF With Complex-Valued Data," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 5, pp. 971–982, May 2013.

[14] "NAO v5, technical documentation," `http://doc.aldebaran.com/2-1/family/robots`, Accessed: 2021-03-01.

[15] M. N. Schmidt, J. Larsen, and F. Hsiao, "Wind Noise Reduction Using Non-Negative Sparse Coding," in *Proc. IEEE Int. Workshop Acoustic Signal Enhancement (IWAENC)*, Thessaloniki, Greece, 2007, pp. 431–436, IEEE.

[16] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for the speech distortion weighted multichannel wiener filter for robust noise reduction," *Speech Communication*, vol. 49, no. 7, pp. 636–656, 2007.

[17] A. Spriet, M. Moonen, and J. Wouters, "Spatially pre-processed speech distortion weighted multi-channel wiener filtering for noise reduction," *Signal Proc.*, vol. 84, no. 12, pp. 2367–2387, 2004.

[18] D. Gouaillier et al., "Mechatronic design of NAO humanoid," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, May 2009, pp. 769–774.

[19] M. Cooke and J. Barker, "An audio-visual corpus for speech perception and automatic speech recognition," *J. Acoustical Society of America*, vol. 120, no. 5, pp. 2421–2424, 2006.

[20] H. Löllmann, H. Barfuss, A. Deleforge, S. Meier, and W. Kellermann, "Challenges in acoustic signal enhancement for human-robot communication," in *Proc. 11. ITG Symp. on Speech Commun.*, Erlangen, Germany, 2014, pp. 1–4, VDE ITG.

[21] C. Févotte, R. Griboval, and E. Vincent, "BSS eval toolbox user guide," Tech. Rep., IRISA, Rennes, France, April 2005.

[22] ITU-T Recommendation P.862.2, "Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs," Recommendation, ITU, Nov. 2007.

GOAL-ORIENTED COMMUNICATION FOR REAL-TIME TRACKING IN AUTONOMOUS SYSTEMS

Nikolaos Pappas[1], *Marios Kountouris*[2]

[1]Department of Science and Technology, Linköping University, Norrköping 60174, Sweden
[2]Communication Systems Department, EURECOM, Sophia-Antipolis 06904, France
Email: nikolaos.pappas@liu.se, marios.kountouris@eurecom.fr

ABSTRACT

Real-time remote tracking using under-sampled and delayed measurements is considered here. We study an autonomous system where a transmitter monitors the evolution of a discrete Markov source and sends status updates to a destination over an unreliable wireless channel. The destination is tasked with real-time source reconstruction for remote actuation. We introduce new goal-oriented sampling and communication policies, which leverage the significance and effectiveness of messages, as a means to generate and transmit only the most "informative" samples for real-time actuation. Our results illustrate that semantics-empowered policies significantly reduce both the real-time reconstruction and the cost of actuation errors, as well as the amount of ineffective updates.

Index Terms— Goal-oriented communication, semantics of information, cost of actuation error, real-time tracking.

1. INTRODUCTION

Networked autonomous systems are ubiquitous in various domains, with applications spanning from swarm robotics and healthcare to autonomous transportation and environmental monitoring. These systems require reliable real-time communication, autonomous interactions, and timely computations. In this context, information is valuable when it is fresh, accurate, and effective. For instance, real-time knowledge of the trajectory and the velocity of a mobile robot is essential in autonomous navigation. Timely and accurate updates are of cardinal importance in mission-critical decision-making. In this setting, real-time tracking and reconstruction of an information source/process from a set of under-sampled and delayed measurements is an important yet challenging problem. Information freshness is assessed by the Age of Information (AoI) [1, 2], i.e., the time elapsed since the newest successfully received update was generated. However, AoI does not take into account the source/process evolution and the application context. Several metrics [3–7] have been employed to address the shortcomings of AoI. A line of work that considers AoI and its variants as a criterion for remote estimation can be found in [8–14]. A new communication paradigm has recently been proposed, which accounts for the *semantics* (i.e., significance, goal-oriented usefulness, and contextual value) of information and leverages the synergy between data processing, information transmission, and signal reconstruction [15–17].

In this work, we consider the problem of real-time tracking and reconstruction of an information source. A transmitter (observer) tracks the state of a Markov source and sends status updates (samples) to a receiver over an unreliable wireless channel. Real-time reconstruction is performed at the destination for the purpose of remote actuation. This fundamental setting could model various real-time applications in autonomous networked systems. We introduce new goal-oriented, semantics-empowered sampling and communication policies, which account for the temporal evolution of the source/process and the semantic and application-dependent value of data being generated and transmitted. Our approach is shown to significantly reduce both reconstruction error and cost of actuation error, as well as the number of uninformative/ineffective samples.

2. SYSTEM MODEL

We consider a time-slotted communication system, in which a monitoring device (transmitter) observes a process X_t and informs a remote actuator (receiver) about its state by sending updates (samples) over a communication channel. This is depicted in Fig. 1. We assume that transmissions take place over a wireless erasure channel. The channel realization h_t is equal to 1 if the packet is successfully decoded at the receiver and 0 otherwise. The success probability is defined as $p_s = \mathbb{P}(h_t = 1)$. Successful/failed transmissions are declared to the transmitter using acknowledgement (ACK)/negative ACK packets, assumed to be delivered instantaneously and error-free. When the transmission fails, the update is discarded (no retransmission). The information source is modeled by a two-state discrete-time Markov chain (DTMC) $\{X_t, t \in \mathbb{Z}_0^+\}$, as-

This work is supported in part by the Swedish Research Council (VR), ELLIIT, and CENIIT. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101003431).

sumed to be ergodic. At each timeslot t, the state X_t of the source can be either 0 or 1. The self-transition probabilities are denoted by $1 \bullet p$ and $1 \bullet q$ for states 0 and 1, respectively. Therefore, $\mathbb{P}\left(X_{t+1}=i|X_t=i\right)=\mathbb{1}\left(i=0\right)\left(1\bullet p\right)+\mathbb{1}\left(i=1\right)\left(1\bullet q\right)$, where $\mathbb{1}(\cdot)$ is the indicator function.

The transmitter is capable of generating update X_t by sampling the source at will. The action of sampling at timeslot t is denoted by α_t^{s}, where $\alpha_t^{\mathrm{s}}=1$ if the source is sampled and $\alpha_t^{\mathrm{s}}=0$ otherwise. The action of transmitting a sample is denoted by $\alpha_t^{\mathrm{tx}}=1$, otherwise the transmitter remains silent ($\alpha_t^{\mathrm{tx}}=0$). The source is reconstructed at the destination based on successfully received status updates. The state of the reconstructed source at timeslot t is denoted as $\hat{X}_t$. A set of operation policies for sampling and transmission are presented in Section 4.

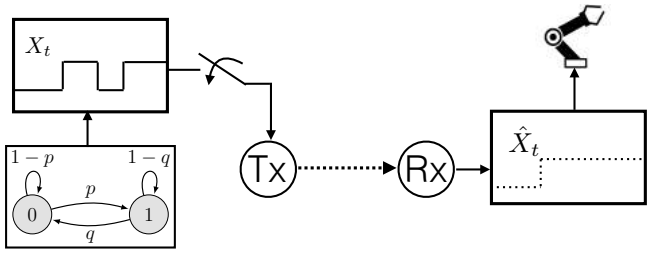


Fig. 1. Real-time tracking of a two-state Markov source.

3. KEY PERFORMANCE METRICS

In this section, we introduce a new, comprehensive system metric, namely the Semantics of Information (SoI), which captures the significance and usefulness of information with respect to the goal of data exchange and the application requirements. We then present two performance metrics used to evaluate the accuracy of real-time reconstruction and its effect on the actuation in our problem under study.

3.1. Semantics-empowered metrics

Let $\mathcal{I}\in\mathbb{R}^m$ denote the vector of m information attributes, which can be decomposed into *innate* (objective) and *contextual* (subjective). Innate are the attributes inherent to information regardless of its use, such as *freshness* or AoI, i.e., $\Delta_t = t\bullet U_t$ where U_t is the generation time of the newest sample that has been delivered at the destination by time instant t, *precision*, and *correctness*. Contextual are attributes that depend on the particular context or application for which information is being used. The most relevant ones are *timeliness*, a function $g:\mathbb{R}_{\geq 0}\to\mathbb{R}$ of AoI, i.e., $g(\Delta_t)$, and *completeness*. Another relevant attribute is *accuracy (distortion)* $\delta:\mathcal{X}\times\mathcal{X}\to\mathbb{R}_{\geq 0}$ where $\mathcal{X}$ is the state space of X_t. For example, $\delta(X_t,\hat{X}_t)=(X_t\bullet\hat{X}_t)^2$. We can also include the notion of *perception* via some divergence or distance function $D(\cdot\|\cdot):\mathcal{D}\times\mathcal{D}\to\mathbb{R}$ between probability distributions defined in the same probability space $\mathcal{D}$ (e.g., Wasserstein, Levenshtein, Hellinger, etc., depending on the application).

Formally, SoI is a composite function $\mathcal{S}_t=\nu(\psi(\mathcal{I}),$ where $\psi:\mathbb{R}^m\to\mathbb{R}^z, m\geq z$ is a nonlinear function and $\nu:\mathbb{R}^z\to\mathbb{R}$ is a context-dependent, cost-aware function that maps qualitative information attributes to their application-dependent value. Below, we provide two such metrics, which are relevant to remote real-time tracking and actuation.

3.2. Real-time reconstruction error

The real-time reconstruction error measures the discrepancy between the original X_t and the reconstructed source $\hat{X}_t$ at timeslot t, i.e., $E_t=\mathbb{1}\left(X_t\neq\hat{X}_t\right)=\left|X_t\bullet\hat{X}_t\right|$. In other words, $\delta(\cdot,\cdot)$ is the $0\bullet 1$ loss function or the Hamming distortion measure. For a two-state DTMC, E_t takes values 0 or 1. Our analysis can easily be generalized to an N-state Markov source ($N<\infty$), in which case the reconstruction error can take any value from 0 to N. The system can be either in an erroneous state ($E_t=1$) or in a synced state ($E_t=0$). The time-averaged reconstruction error is given by

$$\bar{E}=\lim_{T\to\infty}\frac{\sum_{t=1}^{T}E_t}{T}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{1}\left(X_t\neq\hat{X}_t\right). \quad (1)$$

The evolution of the state of the system, E_t, can be described by a Markov Chain as depicted in Fig. 2. The synced state is denoted by 0 ($E_t=0$ at timeslot t), whereas 1 denotes the erroneous state. The one-step transitions probabilities are defined as

$$p_{ji}=\mathbb{P}\left(E_{t+1}=j|E_t=i\right),\forall i,j\in\{0,1\}. \quad (2)$$

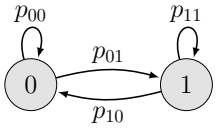


Fig. 2. DTMC describing the state E_t of the system.

We now give general expressions for the transition probabilities. To obtain p_{00} we need to calculate

$$p_{00}=\sum_{i=0}^{1}\mathbb{P}\left(E_{t+1}=0|E_t=0,X_t=i\right)\mathbb{P}\left(X_t=i\right), \quad (3)$$

with $\mathbb{P}\left(X_t=i\right)=\mathbb{1}\left(i=0\right)\frac{q}{p+q}+\mathbb{1}\left(i=1\right)\frac{p}{p+q}$ and

$$\mathbb{P}\left(E_{t+1}=0|E_t=0,X_t=0\right)= \\ =1\bullet p+p\mathbb{P}\left(\alpha_{t+1}^{\mathrm{s}}=1,\alpha_{t+1}^{\mathrm{tx}}=1,h_{t+1}=1\right). \quad (4)$$

In a similar way, we obtain $\mathbb{P}\left(E_{t+1}=0|E_t=0,X_t=1\right)$. We also have

$$\mathbb{P}\left(E_{t+1}=1|E_t=1,X_t=0\right)=(1\bullet p) \\ \times\left[\mathbb{P}\left(\alpha_{t+1}^{\mathrm{s}}=1,\alpha_{t+1}^{\mathrm{tx}}=1,h_{t+1}=0\right)+\mathbb{P}\left(\alpha_{t+1}^{\mathrm{s}}=0\right)\right]. \quad (5)$$

The exact values of transition probabilities depend on the sampling and transmissions actions, dictated by the policies introduced in Section 4. We can also compute the stationary distribution of this two-state DTMC, where π_0 is the probability the system is synced (or the percentage of time the system is synced), and π_1 the probability the system is in an erroneous state.

3.3. Cost of actuation error

The second performance metric is the cost of actuation error, which captures the significance of the error at the point of actuation. Note that some errors may have larger impact than others. At timeslot t, $C_{i,j}$ denotes the cost of being in state i at the original source and in $j \neq i$ at the reconstructed, i.e., $E_t = 1$. Whenever $i = j$, there is no error, and consequently no cost. We consider the general and practically relevant case of non-commutative errors, i.e., $C_{0,1} \neq C_{1,0}$. This means that different erroneous actions may have different cost (penalty) due to different repercussions for the system performance. We assume that C_{ij} does not change over time.[1]

In order to calculate the average cost of actuation error, we can use a two-dimensional Markov chain describing the joint status of the system regarding the current state at the original source and whether the reconstructed source is synced or not. That way, we can obtain expressions for the average real-time reconstruction error and the average cost of actuation error. The latter can be written as

$$\bar{C}_A = \pi_{(0,1)} C_{0,1} + \pi_{(1,0)} C_{1,0} \tag{6}$$

where $\pi_{(i,i+1 \pmod 2)}$ is obtained from the stationary distribution of the two-dimensional DTMC.

Remark. The above Markov chain formulation provides a very general view of the system, which can be used to derive optimal online policies using Markov Decision Processes or Deep Reinforcement Learning.

4. GOAL-ORIENTED SAMPLING AND COMMUNICATION POLICIES

In this section, we propose two semantics-empowered policies of information acquisition (sampling) and transmission for real-time reconstruction of a Markov source with the purpose of actuation. We start by presenting two conventional policies, which are used for comparison purposes. Due to space limitations, expressions only for the last goal-oriented policy are provided; analyzing the other policies involves a modification of the general expressions given in Section 3.

4.1. Uniform

In this baseline policy, sampling is performed periodically, independently of the temporal evolution of the source. Despite being simple and easy to implement, process-agnostic policies could result in missing several state transitions during the time interval between two collected samples. In the case of erasure, the most recently acquired sample is transmitted.

4.2. Age-aware

In this policy, the receiver triggers the acquisition and transmission of a new sample, once the AoI reaches a predefined threshold A_{th}. This can be extended to different AoI thresholds depending on the state [18].

Whenever a transmission fails, the receiver tries to anticipate the update based on the statistics of the source process. In that case, the receiver, given its current state, tries to predict the next state based on the state transition probabilities, assumed to be known (or learned after a period of time). This policy remains source-agnostic regarding the value of information but takes into account the timeliness.

4.3. Change-aware

In this policy, sample generation is triggered at the transmitter whenever a change at the state of the source (with respect to the previous sample) is observed. Consider that for a certain period of time $X_{t+kt} = i, k = 0, 1, \ldots, K$, at the end of which the state changes, i.e., $X_{t(K+1)+1} = j, j \neq i$, and hence the transmitter generates and transmits a new status update sample.

4.4. Semantics-aware

This policy extends the previous one into that the amount of change is not solely measured at the source, but is also tracked by the difference in state between receiver and transmitter. Sample generation is triggered whenever there is discrepancy between X_t and $\hat{X}_t$. Assume that at a given timeslot t, $X_t = \hat{X}_t$. Then, a change in X_t (source's state) occurs in $t+1$ with a given probability, resulting in transmission of a newly acquired sample. If transmission fails, $\hat{X}_{t+1} = \hat{X}_t$. Suppose now that in the next slot, $X_{t+2} = X_t$. Thus, there is no discrepancy between the original and the reconstructed source ($\hat{X}_{t+2} = X_{t+2}$), thus, there is no need for sending an update. Particularizing the general expressions in Section 3, the transition matrix P_E for the DTMC that models the system state is given by

$$\mathrm{P}_E = \begin{bmatrix} 1 \bullet \frac{2pq(1-p_s)}{p+q} & \frac{2pq(1-p_s)}{p+q} \\ p_s + \frac{2pq(1-p_s)}{p+q} & 1 \bullet p_s \bullet \frac{2pq(1-p_s)}{p+q} \end{bmatrix}.$$

Then, we obtain the the probability that the system is in an erroneous (not synced) state

$$\bar{E} = \frac{2pq(1 \bullet p_s)}{p_s(p+q) + 4pq(1 \bullet p_s)}. \tag{7}$$

Note that the previous expression is also the percentage of time that the system is not synced or the time-average reconstruction error.

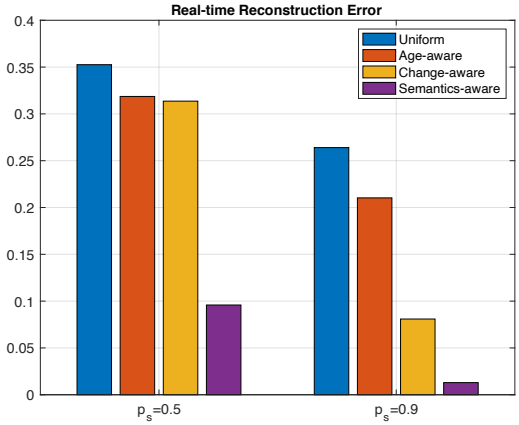
Remark. Evidently, sampling and transmission at every timeslot could provide the best performance for achieving the goal (real-time reconstruction). However, this comes at the

[1] Penalty functions based on non-linear aging [3], where the cost of being in an erroneous state is increasing over time, can be employed.

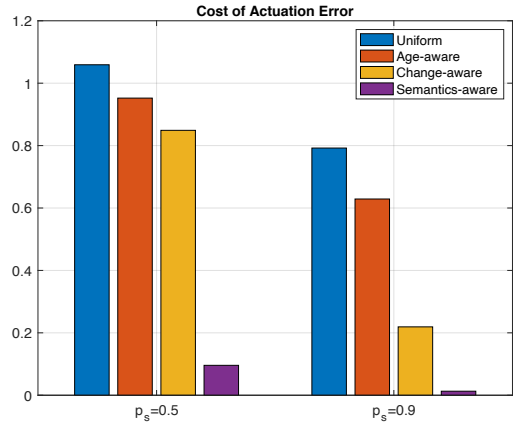
expense of very large number of samples, which are not necessarily useful and which require excessive resources (e.g., energy, network) for their acquisition, transmission, and processing. The proposed semantics-empowered policies reduce or even eliminate the generation of uninformative sample updates, thus improving network resource usage and being scalable.

5. NUMERICAL RESULTS

We evaluate the performance of above policies in terms of average real-time reconstruction error and average cost of actuation error. We consider two scenarios regarding the source variability, the first being when the source is slowly changing ($p = 0.1$, $q = 0.15$), depicted in Fig. 3, and the second being when the source is rapidly changing ($p = 0.2$, $q = 0.7$), depicted in Fig. 4. In addition, regarding the probability of successful transmission we consider two distinct cases, $p_s = 0.5$ and $p_s = 0.9$, to investigate the impact of transmission errors on reconstruction and actuation. In uniform sampling, a sample is acquired every 5 timeslots, and for the age-aware policy we set $A_{\text{th}} = 5$. Additionally, the actuation errors are $C_{0,1} = 5$ and $C_{1,0} = 1$.



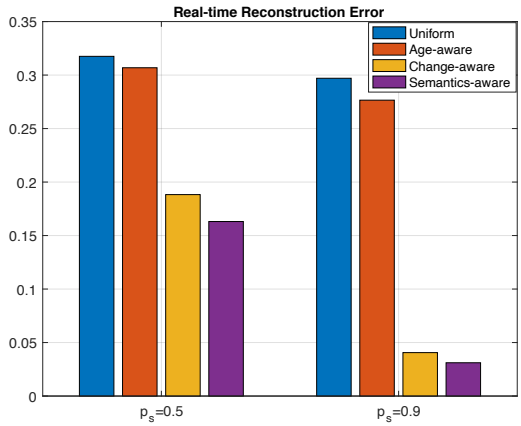
(a) Real-time reconstruction error



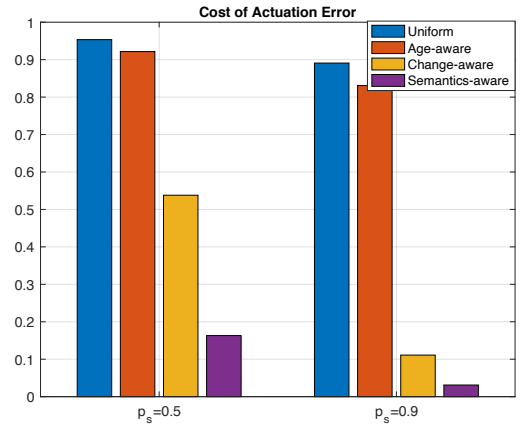
(b) Cost of actuation error

Fig. 3. Slowly varying source with $p = 0.1$ and $q = 0.15$.

For slow varying sources, the semantics-aware policy significantly outperforms the change-aware policy due to the fact that the system manages to rapidly eliminate the discrepancy between the original and the reconstructed state of the source, even in the case of low channel quality ($p_s = 0.5$). On the other hand, for rapidly varying sources, both semantics- and change-aware policies exhibit very good reconstruction error performance, while semantics-aware provides the lowest actuation error *without wasting any resources for transmitting uninformative samples*.



(a) Real-time reconstruction error



(b) Cost of actuation error

Fig. 4. Rapidly changing source with $p = 0.2$ and $q = 0.7$.

6. CONCLUSION

In this work, we showcased the potential of goal-oriented data generation and communication policies in a remote real-time tracking and actuation scenario. Accounting for the semantics of information semantics could lead to significant reduction in task-dependent error metrics and in ineffective traffic volume.

7. REFERENCES

[1] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Foundations and Trends in Networking*, vol. 12, no. 3, 2017.

[2] R. D. Yates, Y. Sun, D. Richard Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of Information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, 2021.

[3] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "The cost of delay in status updates and their value: Non-linear ageing," *IEEE Trans. on Communications*, vol. 68, no. 8, 2020.

[4] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks*, vol. 21, no. 3, 2019.

[5] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, 2020.

[6] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-information vs. value-of-information scheduling for cellular networked control systems," in *10th ACM/IEEE ICCPS*, Apr. 2019.

[7] A. Molin, H. Esen, and K. H. Johansson, "Scheduling networked state estimators based on value of information," *Automatica*, vol. 110, 2019.

[8] T. Z. Ornee and Y. Sun, "Sampling for remote estimation through queues: Age of information and beyond," in *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, 2019.

[9] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the Wiener process for remote estimation over a channel with random delay," *IEEE Transactions on Information Theory*, vol. 66, no. 2, 2020.

[10] C. Kam, S. Kompella, and A. Ephremides, "Age of incorrect information for remote estimation of a binary markov source," in *IEEE INFOCOM Workshops*, 2020.

[11] C. Tsai and C. Wang, "Unifying AoI minimization and remote estimation—optimal sensor/controller coordination with random two-way delay," in *IEEE Conference on Computer Communications (INFOCOM)*, 2020.

[12] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Real-time remote estimation with hybrid arq in wireless networked control," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, 2020.

[13] M. Wang, W. Chen, and A. Ephremides, "Real-time reconstruction of a counting process through first-come-first-serve queue systems," *IEEE Transactions on Information Theory*, vol. 66, no. 7, 2020.

[14] X. Zheng, S. Zhou, and Z. Niu, "Urgency of information for context-aware timely status updates in remote control systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, 2020.

[15] P. Popovski, O. Simeone, F. Boccardi, D. Gündüz, and O. Sahin, "Semantic-effectiveness filtering and control for post-5G wireless connectivity," *Journal of the Indian Institute of Science*, vol. 100, no. 2, 2020.

[16] E. Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," *Computer Networks*, p. 107930, 2021.

[17] M. Kountouris and N. Pappas, "Semantics-empowered communication for networked intelligent systems," *to appear, IEEE Commun. Magazine*, 2021.

[18] G. Stamatakis, N. Pappas, and A. Traganitis, "Control of status updates for energy harvesting devices that monitor processes with alarms," in *IEEE Globecom Workshops*, 2019.

Intelligent Intersection Coordination and Trajectory Optimization for Autonomous Vehicles

Yixiao Zhang, Gang Chen and Tingting Zhang
School of Electronics and Information Engineering,
Harbin Institute of Technology, Shenzhen.
Email: zhangtt@hit.edu.cn

***Abstract*—Since multiple roads merge at intersections, proper coordination for vehicles is of great importance for modern intelligent transportation systems (ITS). In this paper, we try to smartly integrate the infrastructure and vehicle-based planners, to achieve feasible and efficient solutions. In detail, the vehicle reference trajectories can be firstly achieved by the high-level infrastructure-based coordination, which can be formulated as standard quadratic programming (QP) and mixed integer programming (MIP) problems. Due to the possible occurrence of obstacles such as pedestrians, the vehicles are also required to perform low-level ego trajectory optimization based on local observations, which are essentially dynamic programming (DP) and QP problems. Numerical results show that the proposed framework can effectively solve many opening problems in vehicle coordination, such as obstacle avoidance and deadlocks among vehicles.**

I. Introduction

With the rapid urbanization nowadays, more and more vehicles are expected to enter the road infrastructure, which makes intelligent transportation systems challenging and important. More than 40% of traffic accidents, including 20% fatal, occur at road intersections where multiple roads merge [1], [2]. Therefore, careful and efficient maneuver behaviors for vehicles at intersections are key solutions to avoid accidents as well as jams.

Traditional intersection coordination solutions include traffic lights, roundabouts and stop signs, etc. To improve the safety and efficiency, a natural trend is to turn to the dedicated systems for help [3]. Currently, depending on whether there exists a *coordination node*[1], the intersection coordination methods can be categorized into infrastructure-based and vehicle-based strategies. In infrastructure-based strategies, the coordination node will allocate space and time resources for vehicles via the requests sending from vehicles [4]. To guarantee the safety, the area where the collisions may occur is defined as the *critical set*, where only one vehicle is allowed to enter simultaneously. The infrastructure-based coordination problems can be usually formulated as mixed integer programming (MIP) problems [5]–[7].

In vehicle-based strategies, vehicles use their own sensors such as the cameras, radars and LIDARs, for environment awareness, and then plan trajectories accordingly [8]. Parametric curves those subject to vehicle kinematic constraints are mainstream geometric methods to describe trajectories, e.g., line & circle curves [9], polynomial curves [10], Bzier curves [11] and splines [12]. Accordingly, numerical optimization, which can model the vehicle and environment constraints in a mathematic form, is widely used in motion planning. In [13], a trajectory optimization method to generate path and speed profile separately was proposed, which can reduce the complexity of the three-dimensional optimization. Meanwhile, the quadratic programming (QP) is a useful tool to search iteratively for an optimal/sub-optimal solution [14]. However, when obstacles such as pedestrians are involved, the trajectory planning problems become non-convex. Baidu Apollo EM planner thus combined dynamic programming (DP) with QP, to give feasible solutions [15].

However, both kinds of strategies have limitations, especially in complex environments. For infrastructure-based coordination strategies, though they have the potential to improve the safety, fuel efficiency and traffic throughput [16], they only consider the cooperation among vehicles and lack of flexibility when dealing with frequent obstacles, such as pedestrians, bicycles and packed cars. On the other hand, the vehicle-based planning strategies mainly try to detect possible obstacles and improve the passenger comfort via trajectory smoothness. But they are usually less efficient in throughput, and even lead to possible *deadlocks* when the number of vehicles increases.

Aiming at the limitations above, we try to integrate both infrastructure and vehicle-based planners in a proper way. The main contributions of this paper can thus be concluded as follows.

- To avoid collisions and deadlocks among vehicles, as well as improving the system efficiency, a high-level coordination strategy (infrastructure-based) is used to provide feasible *reference trajectories* for all vehicles.
- Based on the reference trajectories, we can further propose a low-level path-speed iterative optimization solution (vehicle-based) for each ego vehicle, to guarantee the safety issue, especially in obstacle-limited scenarios.
- The proposed intersection coordination framework can be formulated as standard optimization problems, such as QP, DP and MIP, which can be effectively solved. Numerical results are also provided to see the performance advantages of the framework.

[1]The coordination node can be physically deployed on the road side unit (RSU) in the intersection.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Fig. 1. Graphic model of the intersection coordination system.

II. System Model

A. Intersection Model

We consider a typical road intersection which consists of R roads with lane width $\mathcal{W}_{\text{lane}}$ (usually we set $R = 4$)[2], as shown in Fig. 1. The intersection area can be spatially divided into three main sections. Arrival Area (AA) is where vehicles arrive at the intersection dynamically. In Buffer Area (BA), each arrived vehicle queues up and waits for the coordination. Conflict Area (CA) represents the intersection area where a potential collision may occur. There also exists a *coordination node* to generate reference trajectories including paths and speed profiles for vehicles.

We use queue length I_r to indicate the number of vehicles waiting in BA at road r. Notations $\mathcal{R} = \{1, 2, \ldots, R\}$ and $\mathcal{I}_r = \{1, \ldots, I_r\}$ represent the sets of roads and vehicles on the r-th road, respectively. Without loss of generality, there are four possible maneuvers when a vehicle passes an intersection, i.e., go straight, turn left, turn right and U turn.

B. Vehicle Kinematic Model

We adopt the kinematic bicycle model [17] for vehicles in this paper. The state parameters of the i-th vehicle on road r can be described by $[x_{i,r}, y_{i,r}, \theta_{i,r}, v_{i,r}]^{\mathrm{T}}$, where $(x_{i,r}, y_{i,r})$, $\theta_{i,r}$ and $v_{i,r}$ are position (in Cartesian Coordinates), heading and velocity, respectively. The control inputs are curvature and acceleration $[\kappa_{i,r}, a_{i,r}]^{\mathrm{T}}$, where $\kappa_{i,r}$ is related to the steering angle ω by $\kappa \propto \tan(\omega)$ and $a_{i,r}$ is related to the throttle and brake of the vehicle. Therefore, the vehicle kinematic state can be uniquely determined by the control inputs and its initial state, i.e.,

$$\begin{bmatrix} \dot{x}_{i,r} \\ \dot{y}_{i,r} \\ \dot{\theta}_{i,r} \\ \dot{v}_{i,r} \end{bmatrix} = v_{i,r} \cdot \begin{bmatrix} \cos(\theta_{i,r}) \\ \sin(\theta_{i,r}) \\ \kappa_{i,r} \\ 0 \end{bmatrix} + a_{i,r} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, r \in \mathcal{R}, i \in \mathcal{I}_r \quad (1)$$

[2]Proposed framework can be extended to irregular intersections.

with $\dot{(\bullet)} := \frac{\partial}{\partial t}(\bullet)$. Two different coordinates including the Cartesian Coordinates and Frenét Frame are adopted for trajectory optimization, which can be described as follows.

1) Trajectory in Cartesian Coordinates: In the Cartesian coordinate system, trajectory functions can be expressed as

$$\begin{cases} \text{Path:} & \begin{cases} x_{i,r} = f_{i,r}(s_{i,r}) \\ y_{i,r} = g_{i,r}(s_{i,r}) \end{cases} \\ \text{Speed Profile:} & s_{i,r} = u_{i,r}(t) \end{cases} \quad (2)$$

with s the arc length of the vehicle along the path. The vehicle controls can be uniquely mapped into the trajectory as follows.

$$[f, g, \arctan(\frac{g'}{f'}), \frac{|f'g'' - f''g'|}{(f'^2 + g'^2)^{3/2}}, u', u'']^{\mathrm{T}} = [x, y, \theta, \kappa, v, a]^{\mathrm{T}} \quad (3)$$

2) Trajectory in Frenét Frame: In the Frenét Frame [18], We use a reference path to reduce the number of functions that describe vehicle path, by setting s and l domains as the longitudinal and lateral distances a vehicle travels along the reference path, respectively. As a consequence, in the new SLT space, trajectory functions can be expressed as

$$\begin{cases} \text{Path:} & l_{i,r} = p_{i,r}(s_{i,r}) \\ \text{Speed Profile:} & s_{i,r} = q_{i,r}(t) \end{cases} \quad (4)$$

From [18], we know that there are transformations

$$[p, q, p', p'', q', q''] \mapsto [x, y, \theta, \kappa, v, a] \quad (5)$$

Therefore, given smooth trajectory functions (2) or (4) for all the vehicles, we can control them via (1), (3) and (5).

C. Structure of the Planning Framework

As shown in Fig. 2, the planning process could be expressed as follows.

- **High-level Planner** (infrastructure-based): The coordination node receives driving maneuvers and initial states of all the vehicles in BA by V2I communication. Then it will generate centralized reference trajectories in Cartesian Coordinates.
- **Low-level Planner** (vehicle-based): Each vehicle first can get its reference trajectory from the coordination node. Based on the reference trajectory, the vehicle will build its practical trajectory using its on-board sensor observations in the Frenét Frame, to avoid unexpected obstacles.

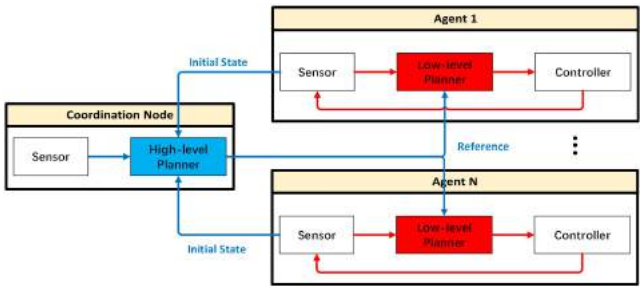


Fig. 2. Structure of the planning framework.

III. High-level Planner: Vehicle Coordination

The high-level planner in the coordination node is to give reference trajectories for all vehicles in BA from all R roads. In this section, we only consider vehicles rather than other obstacles. There are two main parts in high-level planner. Reference path generator (RPG) aims to give vehicles feasible paths for all possible maneuvers, while reference speed profile generator (RSPG) coordinates the speed profiles of vehicles to avoid collisions and improve the traffic performance as well.

A. Reference Path Generator (RPG)

One popular reference path is the simplified line & circle path [9]. However, for practical turning maneuvers, the line & circle paths are too ideal to be realized in the kinematic bicycle model (1). Therefore, we use smooth optimization to generate feasible reference paths based on the line & circle paths, i.e., search for the feasible reference path functions $f(s)$ and $g(s)$[3].

1) Path Parameterization: We segment the total distance of the line & circle path at intervals Δs_{k} and Δs_{a} respectively ($\Delta s_{\mathrm{k}} > \Delta s_{\mathrm{a}}$), for getting $n_{\mathrm{k}}+1$ knots and $n_{\mathrm{a}}+1$ anchor points, which can be described as

$$\begin{cases} \text{Knots} : \{(x_{\mathrm{k},m}, y_{\mathrm{k},m}, s_{\mathrm{k},m}) \quad m = 0,1,...,n_{\mathrm{k}}\} \\ \text{Anchor Points} : \{(x_{\mathrm{a},j}, y_{\mathrm{a},j}, s_{\mathrm{a},j}) \quad j = 0,1,...,n_{\mathrm{a}}\} \end{cases} \tag{6}$$

Then the reference path between every two knots can be parameterized as the two-dimensional quintic polynomial, i.e.,

$$\begin{cases} x = f_m(s) = k_{m0} + k_{m1}(s - s_{\mathrm{k},m}) + ... + k_{m5}(s - s_{\mathrm{k},m})^5 \\ y = g_m(s) = b_{m0} + b_{m1}(s - s_{\mathrm{k},m}) + ... + b_{m5}(s - s_{\mathrm{k},m})^5 \end{cases} \tag{7}$$

with $s \in [s_{\mathrm{k},m}, s_{\mathrm{k},m+1}), m = 0,1,...,n_{\mathrm{k}}-1$. Therefore the optimization problem is transformed into a search for optimal coefficients $\mathbf{P}_{\text{coe}}$.

$$\mathbf{P}_{\text{coe}} = \begin{bmatrix} \mathbf{K}_0^{\mathrm{T}} & \mathbf{K}_1^{\mathrm{T}} \cdots \mathbf{K}_{n_{\mathrm{k}}-1}^{\mathrm{T}} & \mathbf{B}_0^{\mathrm{T}} & \mathbf{B}_1^{\mathrm{T}} \cdots \mathbf{B}_{n_{\mathrm{k}}-1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{8}$$

$$\mathbf{K}_m = \begin{bmatrix} k_{m0} & k_{m1} \cdots k_{m5} \end{bmatrix}^{\mathrm{T}} \tag{9}$$

$$\mathbf{B}_m = \begin{bmatrix} b_{m0} & b_{m1} \cdots b_{m5} \end{bmatrix}^{\mathrm{T}} \tag{10}$$

2) Optimization: We then search for a smooth feasible path near the line & circle path through the following optimization.

$$\mathscr{P}_1 : \min . \sum_{z=2}^{3} w_z^{\mathrm{P}} \left[\int \left(f^{(z)}\right)^2(s)\mathrm{d}s + \int \left(g^{(z)}\right)^2(s)\mathrm{d}s \right] \tag{11}$$

$$\text{s.t.} \quad |f(s_{\mathrm{a},j}) - x_{\mathrm{a},j}| \le \varepsilon \tag{12}$$

$$|g(s_{\mathrm{a},j}) - y_{\mathrm{a},j}| \le \varepsilon \tag{13}$$

$$f_m^{(\gamma)}(s_{\mathrm{k},m+1}) = f_{m+1}^{(\gamma)}(s_{\mathrm{k},m+1}) \tag{14}$$

$$g_m^{(\gamma)}(s_{\mathrm{k},m+1}) = g_{m+1}^{(\gamma)}(s_{\mathrm{k},m+1}) \tag{15}$$

$$\text{Constraints at start and end points} \tag{16}$$

with $m = 0,1,...,n_{\mathrm{k}}-2$; $j = 0,1,...,n_{\mathrm{a}}$ and $\gamma = 0,1,2,3$.

Cost function (11) is to make $\kappa_{i,r}$ continuous and improve the smoothness of the path according to (3). Constraints (12) - (13) limit the search space based on the anchor points by setting ε the maximum allowable offset in both X and Y domains. Thus the behavior of each vehicle could not change heavily. Constraints (14) - (15) ensure that the polynomials are joint at spline knots, by matching the γ-th-order derivative.

Corollary 1: The optimization problem $\mathscr{P}_1$ can be converted to a standard QP problem as follows.

$$\begin{aligned} \mathscr{P}_2 : \min . \quad & \mathbf{P}_{\text{coe}}{}^{\mathbf{T}}\mathbf{H}\mathbf{P}_{\text{coe}} \\ \text{s.t.} \quad & \mathbf{C}_{\text{ineq}}\mathbf{P}_{\text{coe}} \le \mathbf{D}_{\text{ineq}} \\ & \mathbf{C}_{\text{eq}}\mathbf{P}_{\text{coe}} = \mathbf{D}_{\text{eq}} \end{aligned}$$

with constant matrixes $\mathbf{H}$, $\mathbf{C}_{\text{ineq}}$, $\mathbf{C}_{\text{eq}}$, $\mathbf{D}_{\text{ineq}}$ and $\mathbf{D}_{\text{eq}}$.

Proof: See related parts in [15].

Therefore, the coordinate node could generate the reference paths via (7) for all vehicles with 4 maneuvers, using QP.

B. Reference Speed Profile Generator (RSPG)

After the path of each vehicle is predefined by RPG, the RSPG is to coordinate and find feasible speed profile functions $u_{i,r}(t)$ $(r \in \mathcal{R}, i \in \mathcal{I}_r)$ for all vehicles in BA.

For avoiding collisions, we should ensure that only one vehicle can occupy CA simultaneously. Based on the $f_{i,r}(s_{i,r})$, $g_{i,r}(s_{i,r})$ from RPG, the distances when each vehicle entering and leaving CA can be easily gotten and defined as $s_{i,r}^{\mathrm{L}}$ and $s_{i,r}^{\mathrm{H}}$. Therefore, any two vehicles' occupation time in CA should not be overlapped, which can be expressed as

$$\begin{aligned} &[u_{i_1,r_1}^{-1}(s_{i_1,r_1}^{\mathrm{L}}) \quad u_{i_1,r_1}^{-1}(s_{i_1,r_1}^{\mathrm{H}})] \cap \\ &[u_{i_2,r_2}^{-1}(s_{i_2,r_2}^{\mathrm{L}}) \quad u_{i_2,r_2}^{-1}(s_{i_2,r_2}^{\mathrm{H}})] = \varnothing \end{aligned} \tag{17}$$

with $r_1, r_2 \in \mathcal{R}, i_1 \in \mathcal{I}_{r_1}$ and $i_2 \in \mathcal{I}_{r_2}$. An efficient solution is to introduce two integer variables $\alpha_{i,r}, \beta_{i,r} \in \{0,1\}$, which are shown in Table I, to simplify (17) at time t [5].

$$u_{i_1,r_1}(t) \ge s_{i_1,r_1}^{\mathrm{H}} - \alpha_{i_1,r_1} \times M \tag{18}$$

$$u_{i_1,r_1}(t) \le s_{i_1,r_1}^{\mathrm{L}} + \beta_{i_1,r_1} \times M \tag{19}$$

$$u_{i_2,r_2}(t) \ge s_{i_2,r_2}^{\mathrm{H}} - \alpha_{i_2,r_2} \times M \tag{20}$$

$$u_{i_2,r_2}(t) \le s_{i_2,r_2}^{\mathrm{L}} + \beta_{i_2,r_2} \times M \tag{21}$$

$$\alpha_{i_1,r_1} + \alpha_{i_2,r_2} + \beta_{i_1,r_1} + \beta_{i_2,r_2} \le 3 \tag{22}$$

where M is a positive constant large enough. Other constraints for coordination are represented as

$$v_{i,r\,\min} \le u'_{i,r}(t) \le v_{i,r\,\max} \tag{23}$$

$$a_{i,r\,\min} \le u''_{i,r}(t) \le a_{i,r\,\max} \tag{24}$$

$$u'_{i,r}(t) \ge 0 \tag{25}$$

Constraints (23)-(24) are limitations for velocity and acceleration. Constraint (25) avoids backing maneuvers in the scenario.

The cost function of RSPG includes two parts. The first part is for the coordination performance, by comparing the velocity with v_{c}, where v_{c} varies for different performance criteria.

- $v_{\mathrm{c}} = v_{\text{eff}}$. Each vehicle has a velocity v_{eff} for providing the optimal fuel consumption and passenger comfort [19], which can achieve high coordination efficiency.

[3]Here the indexes i and r are omitted.

TABLE I
THE MAPPING BETWEEN BINARY DECISION VARIABLES AND STATE OF THE VEHICLE AT TIME t

Binary decision variables	**State of the i^{th} vehicle at road r**
$\alpha_{i,r}=1, \beta_{i,r}=1$	Inside the CA
$\alpha_{i,r}=0, \beta_{i,r}=1$	Has already passed the CA
$\alpha_{i,r}=1, \beta_{i,r}=0$	Has not yet reached the CA

- $v_{\text{c}} = v_{i,r\,\max}$. If all vehicles can cross the intersection at the highest speed, the intersection throughput can be maximized.

The second part is for keeping $a_{i,r}$ continuous and improving the smoothness of the speed profile.

$$C_{\text{RSPG}} = \underbrace{w_1^{\text{S}} \int \left(v_{\text{c}} - u'_{i,r}(t)\right)^2 \mathrm{d}t}_{\text{Performance}} + \underbrace{\sum_{z=2}^{3} \left[w_z^{\text{S}} \int \left(u_{i,r}^{(z)}\right)^2 (t) \mathrm{d}t \right]}_{\text{Smoothness}} \tag{26}$$

Therefore, the coordination of RSPG can be described as

$$\begin{aligned} \mathscr{P}_3: \ \min. \quad & \sum_{r=1}^{R} \sum_{i=1}^{I_r} C_{\text{RSPG}}(u_{i,r}(t)) \\ \text{s.t.} \quad & (16), (18)-(25) \end{aligned}$$

The MIP problem in $\mathscr{P}_3$ can be solved by many off-the-shelf solvers, such as MOSEK, YALMIP, etc. We can also use an efficient computation offloading method when solving the coordination problem, which can reduce the computation time of the system [20].

IV. Low-level Planner: Vehicle Ego Replanning

In real intersections, in addition to autonomous vehicles, there may exist various obstacles such as pedestrians, bicycles, parked cars. Therefore, we need to consider the trajectory optimization issue for obstacle avoidance based on the RPG/RSPG results generated from the coordination node in Section III. As shown in Fig. 3, we can set the reference path as the center line and build a virtual lane with lane width $\mathcal{W}_{\text{lane}}$. Each vehicle will plan its practical trajectory accordingly.

Furthermore, due to the limited perception range of most vehicle-mounted sensors, we assume a circular perception field for each vehicle. Therefore, the low-level ego replanning strategy on each vehicle should be done iteratively to handle the oncoming obstacles to the perception field.

For each vehicle, firstly, Frenét Frame should be built based on the reference path given from the coordination node. Then low-level planner is used to find feasible trajectory functions $p(s)$, $q(t)$[4], using the path-speed iterative algorithm [13] and expectation maximum (EM) algorithm [15]. The cost functions are given in Table II.

Path planning and speed profile planning are implemented in SL and ST coordinates respectively. For each planning, we map the obstacles and build lattice on the coordinates. All the

[4]Here the indexes i and r are omitted.

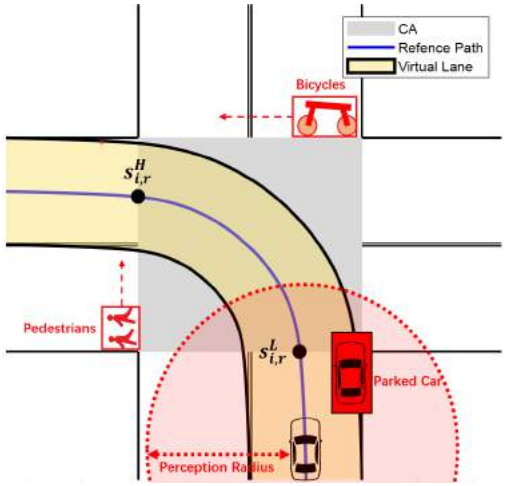


Fig. 3. Virtual lane for low-level planner (e.g. turning left maneuver)

static and dynamic obstacles in the perception field are under consideration. DP can be used to get the set of lattice points with a minimum $C_{\text{s}} + C_{\text{b}} + C_{\text{c}}$, which can be described by

$$\text{Path:} \quad \{(s^{\text{P}}_{\text{k},m}, l^{\text{P}}_{\text{k},m}) \ \ m = 0, 1, ..., n^{\text{P}}_{\text{k}}\} \tag{27}$$

$$\text{Speed Profile:} \quad \{(t^{\text{S}}_{\text{k},m}, s^{\text{S}}_{\text{k},m}) \ \ m = 0, 1, ..., n^{\text{S}}_{\text{k}}\} \tag{28}$$

Convex hull $[\varepsilon_{\text{low}}, \varepsilon_{\text{high}}]$ can be obtained based on (27) and (28). Therefore, the path planning problem can be transformed into the path optimization in the convex region, i.e.

$$\begin{aligned} \mathscr{P}_4: \ \min. \quad & C^{\text{P}}_{\text{s}}(p(s)) + C^{\text{P}}_{\text{b}}(p(s)) \\ \text{s.t.} \quad & (14), (16) \\ & \varepsilon^{\text{P}}_{\text{low}}(s) \le p(s) + \delta p'(s) \le \varepsilon^{\text{P}}_{\text{high}}(s) \end{aligned} \tag{29}$$

where we replace f and s_{k} with p and s^{P}_{k} in (14). Constraint (29) is to keep the path within the convex hull, $\delta p'(s)$ is the offset considering vehicle heading [15].

The speed profile optimization can be given by

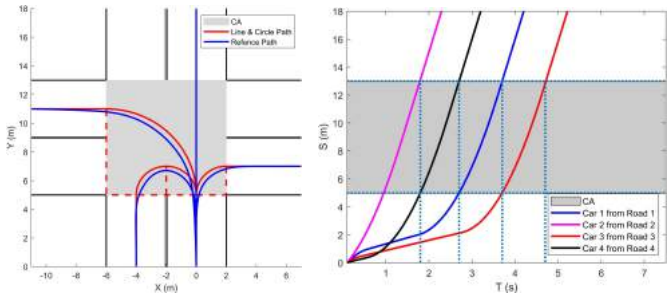
$$\begin{aligned} \mathscr{P}_5: \ \min. \quad & C^{\text{S}}_{\text{s}}(q(t)) + C^{\text{S}}_{\text{b}}(q(t)) \\ \text{s.t.} \quad & (14), (16), (23)-(25) \\ & \varepsilon^{\text{S}}_{\text{low}}(t) \le q(t) \le \varepsilon^{\text{S}}_{\text{high}}(t) \qquad (30) \\ & q(u^{-1}_{i,r}(s^{\text{L}}_{i,r})) \le s^{\text{L}}_{i,r} \qquad (31) \\ & q(u^{-1}_{i,r}(s^{\text{H}}_{i,r})) \ge s^{\text{H}}_{i,r} \qquad (32) \end{aligned}$$

where f, s_{k} are replaced by q, t^{S}_{k} in (14) and $u_{i,r}$ is replaced by q in (23) - (25). Constraint (30) keeps the speed profile within the convex hull. Constraints (31) - (32) are to ensure that the low-level planner still follows the non-collision instruction in the high-level coordination solution. So vehicle should not enter CA before $u^{-1}_{i,r}(s^{\text{L}}_{i,r})$ and leave CA after $u^{-1}_{i,r}(s^{\text{H}}_{i,r})$.

Similarly, $\mathscr{P}_4$ and $\mathscr{P}_5$ can be solved by spline-based QP to obtain optimal $p(s)$ and $q(t)$. Subsequently, we can combine them to get a feasible practical trajectory. If both $\mathscr{P}_4$ and $\mathscr{P}_5$ cannot find the solution, constraint (32) in $\mathscr{P}_5$ should be removed. The new leaving time $q^{-1}(s^{\text{H}}_{i,r})$ will be sent back to the coordination node, and it will replan the reference trajectories for the rest vehicles.

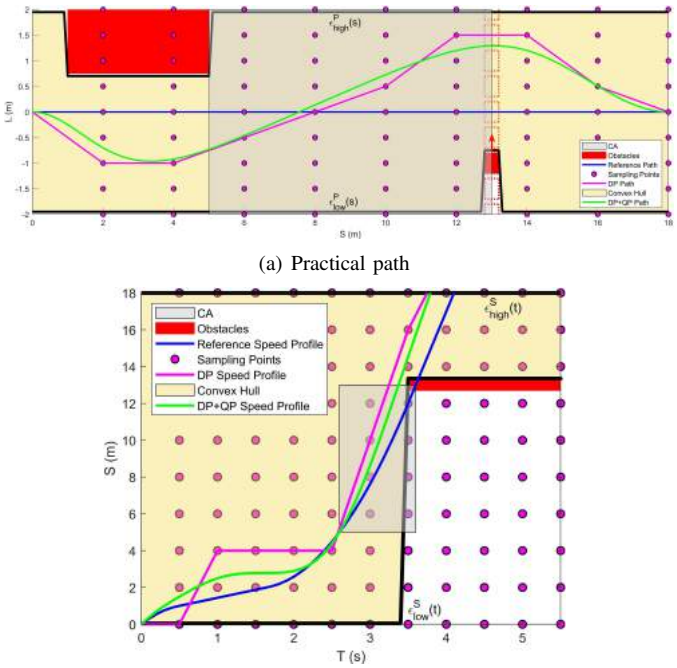
TABLE II
COST FUNCTIONS FOR LOW-LEVEL PLANNER

Impact	Path planning	Speed Profile Planning	Description
smoothness	$C_{\mathrm{s}}^{\mathrm{P}}=\sum_{z=2}^{3}\left[w_z^{\mathrm{P}}\int\left(p^{(z)}\right)^2(s)\mathrm{d}s\right]$	$C_{\mathrm{s}}^{\mathrm{S}}=\sum_{z=2}^{3}\left[w_z^{\mathrm{S}}\int\left(q^{(z)}\right)^2(t)\mathrm{d}t\right]$	improve the passenger comfort
vehicle behavior	$C_{\mathrm{b}}^{\mathrm{P}}=w_4^{\mathrm{P}}\int p^2(s)\mathrm{d}s$	$C_{\mathrm{b}}^{\mathrm{S}}=w_4^{\mathrm{S}}\int\left(q(t)-u_{i,r}(t)\right)^2\mathrm{d}t$	the difference with reference trajectory
collision avoidance	$C_{\mathrm{c}}^{\mathrm{P}}=\sum_{j=1}^{N_{\mathrm{ob}}}\begin{cases}\mathrm{Inf} & d_j^{\mathrm{s}}<d_{j,\min}^{\mathrm{s}}\ \&\ d_j^{\mathrm{l}}<d_{j,\min}^{\mathrm{l}}\\ 0 & \text{else}\end{cases}$	$C_{\mathrm{c}}^{\mathrm{S}}=\sum_{j=1}^{N_{\mathrm{ob}}}\begin{cases}\mathrm{Inf} & d_j^{\mathrm{s}}<d_{j,\min}^{\mathrm{s}}\\ 0 & \text{else}\end{cases}$	compare the distances d^{s},d^{l} and safe distances $d_{\min}^{\mathrm{s}}$,$d_{\min}^{\mathrm{l}}$ with N_{ob} obstacles [21]



(a) Reference paths in Road 1 (b) Reference speed profiles

Fig. 4. Results in high-level planner.



(a) Practical path

(b) Practical speed profile

Fig. 5. Results in low-level planner.

V. NUMERICAL RESULTS

In this part, four roads are considered in a typical traffic intersection, the same as Fig. 1. The simulation settings of main parameters are shown in Table III.

For the high-level planner, Fig. 4(a) shows the reference paths for all 4 driving maneuvers in Road 1, which are also the same in Road $2, 3, 4$. Comparing to simplified line & circle paths, the smooth optimization improves the feasibility and smoothness of the reference paths. In Fig. 4(b), we show the coordination of the vehicle reference speed profiles. All 4 vehicles from different roads are allowed to pass the intersection consecutively. At any time, only one vehicle occupies CA (grey area) and no collision occurs during the coordination.

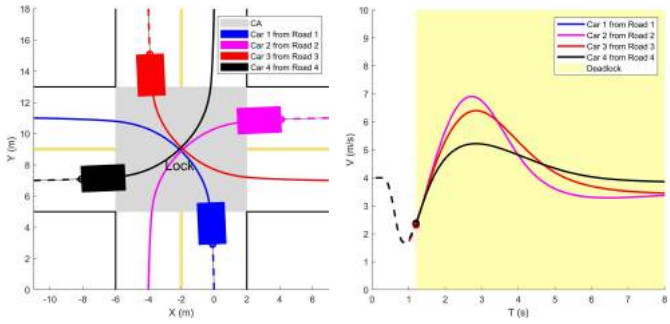
For the low-level path and speed profile planning in Fig. 5(a) and Fig. 5(b), we can get the optimal sampling set (purple line) and convex hull (yellow area) through DP. Then QP is used to obtain practical path and speed profile (green line). We show that the vehicle can avoid collisions with obstacles by nudging and accelerating. In addition, Fig. 5(b) shows that the vehicle's occupation time in CA is within $[u_{i,r}^{-1}(s_{i,r}^{\mathrm{L}})\ u_{i,r}^{-1}(s_{i,r}^{\mathrm{H}})]$, which satisfy the result in high-level coordination.

As mentioned before, combining with the high-level planner, our planning framework can avoid deadlock situations. Using only the low-level planning strategy in Section IV, as shown in Fig. 6(b), 4 vehicles always make the same decision to avoid collisions. Firstly, they all decelerate, and then they all plan to accelerate after perceiving the trends of each other. Therefore, 4 vehicles are locked in CA. By contrast, our proposed framework first determines the crossing order, which shows better performance w.r.t the increasing vehicle number.

We compare the proposed framework with a framework combining traffic lights and low-level planner [22]. In Fig. 7, we can see that the coordination time increases in both frameworks with respect to the vehicle number ($\sum_{r=1}^{R} I_r$). Furthermore, we can see that there are some obvious changes in the traffic light strategy, which is related to the constant change interval of the traffic signal. However, in the proposed framework, the queue lengths I_r of vehicles are considered in the coordination, which leads to a higher throughput.

TABLE III
CALIBRATION OF MAIN PARAMETERS

Parameters	Value	Parameters	Value
$v_{\min}, v_{\max}$	-20m/s, 20m/s	$\Delta s_{\mathrm{k}}, \Delta s_{\mathrm{a}}$	2m, 0.1m
$a_{\min}, a_{\max}$	$-5\mathrm{m/s}^2, 5\mathrm{m/s}^2$	$w_2^{\mathrm{P}}, w_3^{\mathrm{P}}, w_4^{\mathrm{P}}$	$10^4, 10^4, 1$
$w_1^{\mathrm{S}}, w_2^{\mathrm{S}}, w_3^{\mathrm{S}}, w_4^{\mathrm{S}}$	$10, 10^4, 10^4, 1$	ε	0.5m
Perception Radius	15m	v_{eff}	10m/s



(a) Vehicle Paths (b) Vehicle Velocities

Fig. 6. Deadlock for only using low-level planning strategy.

We also compare the proposed framework with a framework combining the simple, sub-optimal *uniform coordination* strategy and low-level planner. In the uniform coordination, vehicles enter the intersection in a constant sequence based on the road number. Fig. 7 shows that our proposed framework has a better throughput performance.

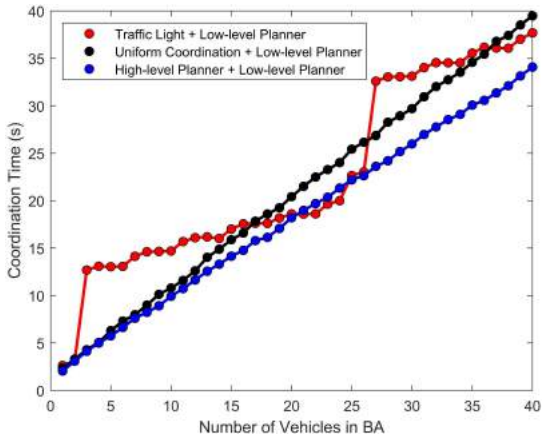


Fig. 7. Coordination time comparisons of different strategies

VI. Conclusion

In this paper, we mainly deal with the trajectory planning problem for autonomous vehicles at road intersections. Due to the possible interactions among vehicles and obstacles, we mainly try to provide a high-level and low-level integrated coordination framework. According to the numerical results, the proposed coordination framework can effectively improve the traffic throughput, without any collisions and deadlocks. This coordination framework, as well as the optimization algorithms, are of importance for the development of autonomous vehicles and future intelligent transportation systems.

VII. Acknowledgment

This work was supported in part by the Natural Science Foundation of China under Grant No. 61771159 and 91638204, and Shenzhen Fundamental Research Project under Grant No. JCYJ20190806143212658.

References

[1] E. H. Choi, *Crash Factors in Intersection-Related Crashes: An On-Scene Perspective*. National Highway Traffic Safety Administration (NHTSA), 2010.

[2] 2010 Motor Vehicle Crashes: Overview. [Online]. Available: {https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811552}

[3] E. Namazi, J. Li, and C. Lu, "Intelligent intersection management systems considering autonomous vehicles: A systematic literature review," *IEEE Access*, vol. 7, pp. 91 946–91 965, 2019.

[4] R. Hult, G. R. Campos, E. Steinmetz *et al.*, "Coordination of cooperative autonomous vehicles: Toward safer and more efficient road transportation," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 74–84, 2016.

[5] R. Hult, G. R. Campos, P. Falcone *et al.*, "An approximate solution to the optimal coordination problem for autonomous vehicles at intersections," in *2015 American Control Conference (ACC)*, 2015, pp. 763–768.

[6] M. Wang, T. Zhang, L. Gao *et al.*, "High throughput dynamic vehicle coordination for intersection ground traffic," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, 2018, pp. 1–6.

[7] C. Liu, Y. Zhang, T. Zhang *et al.*, "High throughput vehicle coordination strategies at road intersections," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14 341–14 354, 2020.

[8] L. Claussmann, M. Revilloud, D. Gruyer *et al.*, "A review of motion planning for highway autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1826–1848, 2020.

[9] J. Horst and A. Barbera, "Trajectory generation for an on-road autonomous vehicle," in *Unmanned Systems Technology VIII*, G. R. Gerhart, C. M. Shoemaker, and D. W. Gage, Eds., vol. 6230, International Society for Optics and Photonics. SPIE, 2006, pp. 866 – 877. [Online]. Available: https://doi.org/10.1117/12.663643

[10] H. Tehrani, Q. Huy Do, M. Egawa *et al.*, "General behavior and motion model for automated lane change," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, 2015, pp. 1154–1159.

[11] F. Garrido, D. González, V. Milanés *et al.*, "Real-time planning for adjacent consecutive intersections," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 1108–1113.

[12] J. Ziegler and C. Stiller, "Spatiotemporal state lattices for fast trajectory planning in dynamic on-road driving scenarios," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 1879–1884.

[13] W. Xu, J. Wei, J. M. Dolan *et al.*, "A real-time motion planner with trajectory optimization for autonomous vehicles," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 2061–2067.

[14] Y. Meng, Y. Wu, Q. Gu *et al.*, "A decoupled trajectory planning framework based on the integration of lattice searching and convex optimization," *IEEE Access*, vol. 7, pp. 130 530–130 551, 2019.

[15] H. Fan, F. Zhu, C. Liu *et al.*, "Baidu apollo em motion planner," *arXiv:1807.08048*, 2018.

[16] Y. Feng, D. He, and Y. Guan, "Composite platoon trajectory planning strategy for intersection throughput maximization," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 6305–6319, 2019.

[17] J. Kong, M. Pfeiffer, G. Schildbach *et al.*, "Kinematic and dynamic vehicle models for autonomous driving control design," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, June 2015, pp. 1094–1099.

[18] M. Werling, J. Ziegler, S. Kammel *et al.*, "Optimal trajectory generation for dynamic street scenarios in a frenét frame," in *2010 IEEE International Conference on Robotics and Automation*, 2010, pp. 987–993.

[19] R. Hult, G. R. Campos, E. Steinmetz *et al.*, "Coordination of cooperative autonomous vehicles: Toward safer and more efficient road transportation," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 74–84, 2016.

[20] Y. Mo, M. Wang, T. Zhang *et al.*, "Intelligent offloading strategies for high throughput traffic intersection coordination," in *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, 2019, pp. 1–6.

[21] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," *arXiv:1708.06374*, 2017.

[22] D. Zhou, Z. Ma, and J. Sun, "Autonomous vehicles' turning motion planning for conflict areas at mixed-flow intersections," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 2, pp. 204–216, 2020.

SEMANTIC IMAGE SEGMENTATION GUIDED BY SCENE GEOMETRY

Sotirios Papadopoulos, Ioannis Mademlis, Ioannis Pitas

Department of Informatics
Aristotle University of Thessaloniki
Thessaloniki, Greece

ABSTRACT

Semantic image segmentation is an important functionality in various applications, such as robotic vision for autonomous cars, drones, etc. Modern Convolutional Neural Networks (CNNs) process input RGB images and predict per-pixel semantic classes. Depth maps have been successfully utilized to increase accuracy over RGB-only input. They can be used as an additional input channel complementing the RGB image, or they may be estimated by an extra neural branch under a multitask training setting. Contrary to these approaches, in this paper we explore a novel regularizer that penalizes differences between semantic and self-supervised depth predictions on presumed object boundaries during CNN training. The proposed method does not resort to multitask training (which may require a more complex CNN backbone to avoid underfitting), does not rely on RGB-D or stereoscopic 3D training data and does not require known or estimated depth maps during inference. Quantitative evaluation on a public scene parsing video dataset for autonomous driving indicates enhanced semantic segmentation accuracy with zero inference runtime overhead.

***Index Terms*—** semantic segmentation, depth estimation, scene geometry, computer vision

1. INTRODUCTION

Semantic image segmentation is one of the most essential scene understanding tasks in modern computer vision, mainly due to its critical importance for autonomous systems, robots and vehicles [1, 2, 3]. It consists in classifying each input image pixel into one amongst a set of prespecified object classes. Convolutional Neural Networks (CNNs) have been the state-of-the-art in similar perception tasks for a long time now. Traditionally, single-view RGB footage has been considered as an adequate input modality for CNNs to successfully perform semantic segmentation. However this is not always the case, since in certain scenarios individual RGB images fail to provide sufficient class-distinctive hints.

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871479 (AERIAL-CORE).

Scene/object geometry has been long known to provide insightful information on several computer vision tasks. Geometry can be described by various formats, such as depth maps, since it can provide cues about shape, texture and distance from the image plane. However, the acquisition of depth maps when constructing an application-specific dataset (e.g., by a depth camera, a LIDAR sensor, etc.) can be a cumbersome task. Typically, high-accuracy depth sensors are expensive and their outputs need heavy post-processing. To alleviate this, a lot of unsupervised/self-supervised depth estimation CNNs have been recently proposed [2]. Such methods learn to infer depth from monocular RGB images by depth supervision [4], by stereo parallax estimation [5], or, as of lately, via monocular video sequences under a neural Structure-from-Motion (SfM) paradigm [6].

The trivial way to exploit this for improving semantic segmentation would be an inference-time two-stage, multimodal approach: estimate a depth map per image/video frame, pair it with corresponding RGB data and feed to it to a neural segmentor pre-trained on RGB-D inputs. Thus, maximum information is extracted from the input image in a pre-processing step performed on-the-fly, facilitating the succeeding CNN in its semantic segmentation task. However, this comes at a significant runtime penalty during model deployment, demanding two different CNNs to be executed in series. An alternative approach would be to train a multitask CNN that concurrently performs both semantic segmentation and depth map estimation from RGB input, using two task-specific neural heads and a common backbone CNN for feature extraction, but this increases training difficulty and requires more complex, slower CNN architectures, able to handle both tasks simultaneously.

This paper presents a novel regularizer that utilizes neurally-estimated (i.e., without requiring any dedicated sensors) depth maps, only while training a conventional semantic segmentation CNN, using regular RGB video data for training (not RGB-D or stereoscopic 3D) and without resorting to a multitask setting (which could demand higher model complexity to avoid underfitting). Thus, the proposed regularizer exploits the depth map modality for increasing segmentation accuracy, without imposing *any* runtime overhead during model deployment/inference and without relying on special

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

input data types at *any* stage. It operates by penalizing differences between semantic and depth predictions on presumed object boundaries. Depth maps are estimated using a separate neural branch, pretrained on regular RGB videos in a self-supervised manner and totally independent from the main segmentation CNN. After training is complete, the latter one can be employed alone, for processing individual, previously unseen single-view RGB images, without relying on dedicated depth sensors or separate geometry estimation CNNs.

Quantitative evaluation of the presented method on a public, scene parsing video dataset for autonomous driving yielded favourable results, in comparison both to the baseline semantic segmentation CNN and to competing methods.

2. RELATED WORK

A vast amount of previous work [7, 8] treats depth as a given input modality to perform computer vision tasks. However these approaches require scene geometry to be known (e.g., by relying on RGB-D sensors), therefore they are not directly related to this paper.

In a number of more relevant cases, depth maps are estimated from RGB input along with semantic segmentation maps, so as to increase accuracy, under a multitask training setting and, typically, with supervised depth estimation (i.e., ground-truth RGB-D data are required during training) [9, 10]. However, certain algorithms falling under this category employ semantic segmentation for extracting improved depth maps, instead of the reverse, and rely on self-supervised depth estimation using stereoscopic 3D images, instead of RGB-D data (e.g., [11, 12, 13]). In [13], enhanced consistency between the two tasks is achieved by inserting a *Cross-Domain Discontinuity* loss term during training, based on the observation that depth discontinuities are likely to co-occur with semantic boundaries. This term detects discontinuities between semantic labels by computing the sign of the absolute value of the gradients in the semantic map. The underlying intuition is that there should be a gradient peak between neighboring pixels belonging to different classes. [12] is a different multitask architecture for semantic image segmentation that is also trained with a similar smoothness loss term. Unlike [13], where the Cross-Domain Discontinuity Term enforces smoothness on the depth values within each ground-truth segmentation mask (thus no error gradients propagate through the segmentation decoder), the regularizer proposed in [12] is computed based on the segmentation branch output. The multitask network shares both the encoder and the decoder, differentiating only in the prediction heads. The decoder is given a task identity signal to predict features for either task.

Certain similar methods employ a pretrained semantic segmentation network to improve depth estimation accuracy, instead of joint multitask training. In two-stage approach [11], depth maps are estimated from stereoscopic 3D image pairs and, subsequently, depth borders are optimized using prior-predicted semantic borders. Then, the method explicitly morphs the predicted depth maps so that depth edges coincide with semantic edges, while finally using this improved depth information as a supervision signal. [14] uses a pretrained segmentor's features to guide a self-supervised depth estimation network's decoder via pixel-adaptive convolutions. Depth estimation relies on video data and on a SfM training loss function.

Few papers report semantic segmentation performance gains by exploiting self-supervised depth estimation. [15] shows that the encoder can have a better weight initialization than simply pretraining on the ImageNet dataset for whole-image classification, by pretraining on automatically computed relative depth derived from self-supervised optical flow; thus the method employs a two-stage training process requiring video data. On the contrary, most other similar approaches rely on a multitask training setting. A number of these multitask methods need stereoscopic 3D training data, such as [16], which trains a multitask network for semantic segmentation, self-supervised depth estimation and image colorization to enhance semantic segmentation performance. On the other hand, [17] estimates depth in a self-supervised manner from regular videos and trains the CNN under a multitask setting with task-specific decoders, achieving a substantial performance increase. [18] also leverages multitask training and self-supervised monocular depth estimation from monocular videos to improve semantic segmentation performance, but it is designed for the special case of semi-supervised learning; thus, it is not directly related to this paper.

Focusing only on papers most similar to ours, the regularizers presented in both [13] and [12] are used mainly to guide depth estimation under a multitask training setting. [16] also employs a multitask architecture, to improve semantic segmentation by exploiting depth estimation. In comparison, the regularizer proposed in this paper is employed for optimizing fully supervised semantic segmentation, without resorting to a multitask architecture or training. Thus, from a different perspective and at a high level of abstraction, the proposed method can be broadly seen as the inverse of [11], which is designed for depth estimation.

3. GEOMETRY-GUIDED SEMANTIC SEGMENTATION

The proposed method does not require RGB-D or stereoscopic 3D training data (which may be difficult to acquire for specific applications), does not impose any runtime overhead during inference on the trained model (as the naive inference-time two-stage approach does) and requires no architectural modifications to the semantic segmentation CNN for facilitating multitask training without underfitting. It consists in: a) self-supervised pretraining of a separate depth map estimation CNN branch and, b) subsequently, training in a regular

manner any conventional semantic segmentation CNN with an additional regularizing loss term, i.e., the proposed *holistic consistency* loss. The latter one is computed at each training iteration using the outputs of the segmentor and of the pretrained depth estimation branch.

The underlying intuitive observation was that semantic objects tend to stand out in depth maps, leading to co-occurence of image gradients in the two tasks. The proposed loss term penalizes semantic map edges that are absent from the spatially corresponding region of the respective depth map, since the target is to enhance semantic segmentation accuracy using depth and not vice versa. As a result, during training, the segmentor is discouraged from outputting semantic shapes that do not conform to scene geometry. Thus, depth information is implicitly integrated while the CNN model is being optimized, but, subsequently, no depth inputs or depth estimation neural branches are required during inference. It is clear that the depth estimation branch can be totally omitted in model deployment, since it is only required during training for computing the proposed regularizer.

3.1. Notation

- $C \in \mathbb{N}$: the number of semantic classes.
- $N_1, N_2 \in \mathbb{N}$: the image spatial dimensions (in pixels).
- $\{A(i,j)\}_{\substack{1 \le i \le N_1, \\ 1 \le j \le N_2}}$: A matrix $\mathbf{A} \in \mathbb{R}^{N_1 \times N_2}$, composed of entries A_{ij}, $1 \le i \le N_1, 1 \le j \le N_2$.
- $\mathbf{S} \in \mathbb{R}^{N_1 \times N_2 \times C}$: the estimated segmentation map. It is a tensor, with each of its C 2D channels being class probability heat maps.
- $\mathbf{D} \in \mathbb{R}^{N_1 \times N_2}$: the estimated depth map. It is a matrix containing the normalized distance of each depicted 3D point from the image plane.
- $mean(\mathbf{A})$: the mean over all entries of matrix $\mathbf{A}$.
- $max(\mathbf{a})$: the maximum over all entries of vector $\mathbf{a}$.
- x, y: the spatial image axes.

3.2. Per-class consistency loss

Here we introduce a preliminary, more involved variant of the proposed holistic consistency loss, in order to facilitate understanding. The *per-class consistency* loss term first computes the degree of consistency between the semantic heat map edges within each channel of $\mathbf{S}$ and the depth map edges at class level. Finally, this quantity is summed up over all classes:

$$L_p = \sum_{c=1}^{C} mean(\{|\frac{dS}{dx}(i,j,c)| \cdot e^{-|\frac{dD}{dx}(i,j)|}\}_{\substack{1 \le i \le N_1, \\ 1 \le j \le N_2}}) + mean(\{|\frac{dS}{dy}(i,j,c)| \cdot e^{-|\frac{dD}{dy}(i,j)|}\}_{\substack{1 \le i \le N_1, \\ 1 \le j \le N_2}}) \quad (1)$$

This formula is based on a simple dissimilarity metric of the form:

$$w(a,b) = |a| \cdot e^{-|b|}. \quad (2)$$

In our case, a/b is semantic/depth edge intensity at a specific image pixel, respectively. In regions with intense depth edges $\lim_{b \to \pm\infty} e^{-|b|} = 0$, therefore semantic edges are not discouraged if depth edges are present. However, in absence of depth edges $\lim_{b \to 0} e^{-|b|} = 1$, so the per-class consistency loss is positive ($w(a,b) > 0$) if, simultaneously, the spatially coinciding semantic map region has a non-zero gradient ($a > 0$).

3.3. Holistic consistency loss

The proposed holistic consistency loss term L_h is a simplification of the per-class consistency loss, where, instead of class-wise edge comparison, the global predicted semantic boundaries are used. These boundaries can be formed by choosing the maximum value among all class semantic edges, for each pixel:

$$L_h = mean(\{S'_x(i,j) \cdot e^{-|\frac{dD}{dx}(i,j)|}\}_{\substack{1 \le i \le N_1, \\ 1 \le j \le N_2}}) + mean(\{S'_y(i,j) \cdot e^{-|\frac{dD}{dy}(i,j)|}\}_{\substack{1 \le i \le N_1, \\ 1 \le j \le N_2}}), \quad (3)$$

where $\mathbf{S}'_k = \{max(|\frac{d\mathbf{S}}{dk}(i,j)|)\}_{\substack{1 \le i \le N_1, \\ 1 \le j \le N_2}}$.

Evidently, the proposed holistic consistency variant is more computationally efficient.

4. EXPERIMENTAL EVALUATION

To assess the proposed L_h method, a popular U-net [6] with a ResNet-50 backbone CNN, pretrained on consecutive video frame pairs, was selected as the depth estimation neural branch, since it does not rely on stereoscopic 3D input. A popular, fast semantic segmentation CNN was selected as the main neural branch [20], serving as our baseline algorithm. The "road01" subset of the Apolloscape dataset [19] was employed for evaluation purposes; to the best of our knowledge, it is the only publicly available and sufficiently large video segmentation dataset (video data are required by [6]). An example video frame is shown in Figure 1. In all cases the original image resolution of 3384×2710 was reduced to 832×256, during both training and testing.

The proposed regularizer L_h was evaluated by comparing the semantic segmentation performance of [20], trained with L_h, against the baseline [20], trained without L_h. Additionally, three properly adapted, recent, competing methods were also evaluated: [17], [12] and [16]. First, the state-of-the-art multitask architecture described in [17] was implemented and the hyperparameter values reported in the paper were further tuned by us for fair comparison. Alternatively, the consistency loss term proposed in [12], which is reminiscent of ours, was implemented on top of [20], instead of L_h, and evaluated under two setups: a) training with a pretrained depth

(a) (b) (c)

Fig. 1: (a) Input image [19], (b) estimated depth map [6], (c) estimated semantic segmentation map (ours))

Table 1: Evaluation results on the Apolloscape dataset, employing a CNN backbone pretrained on ImageNet for image classification. The baseline semantic segmentation branch is [20] and the depth estimation branch is [6]. Methods reported as "pretrained" use a pretrained depth estimation branch with frozen parameters, while methods reported as "multitask" jointly train the two branches. Reported inference time per video frame is an average over the test set.

Method	Mean IoU	Inference runtime (msec)
Baseline (no depth)	39.557%	6.2
[17] (multitask)	34.318%	6.4
Baseline + [16] (multitask)	37.683%	8.3
Baseline + [12] regularizer (pretrained)	39.610%	6.2
Baseline + [12] regularizer (multitask)	38.153%	9
Baseline + L_h (pretrained, proposed)	**40.597**%	**6.2**

network branch (as in the proposed method), and b) training it under a multitask setting, where the semantic segmentation and the depth estimation tasks are learnt simultaneously (as in the original [12]). Finally, the multitask method [16] (without a colorization decoder) was also implemented and plugged on top of [20]. Note that the original [12] and [16] algorithms utilize stereoscopic 3D images during multitask training, thus requiring special datasets. Therefore, for fair comparison, we adapted them to our setting of self-supervised depth estimation from regular, monocular, RGB video using [6], which is arguably a more difficult task. Vanilla [17] already relies on self-supervised depth estimation from RGB video, as is the case with the proposed method.

The evaluation results are depicted in Table 1. Training with the competing loss term from [12] leads only to marginal performance increases over the baseline, since this term mainly improves estimated depth map accuracy, which in turn may offer better scene geometry insights for semantic segmentation. Moreover, in the cases of [17], [12] and [16], multitask training for joint semantic segmentation and self-supervised depth estimation from video does not improve segmentation performance. If the backbone CNN's complexity is not increased, it is a particularly difficult task for multitask learning to handle, compared to using a pretrained depth estimator, thus leading to underfitting. This highlights the advantage of the proposed method in comparison to competing multitask training approaches, when low model complexity (and, thus, low runtime inference requirements) is an important consideration, as in autonomous systems and robotics applications.

Overall, training with the proposed L_h gives a boost of about 1% in the common mIoU metric over the baseline, while also surpassing [17] and the adapted versions of [12] and [16]. The main advantage of the proposed method is its ability to enhance semantic segmentation performance with zero runtime inference overhead, without requiring a complex CNN model, special ground-truth training data (RGB-D, stereoscopic 3D) or special sensors during deployment. Thus, it is most suited to embedded applications such as autonomous systems, robots, vehicles, etc.

5. CONCLUSIONS

This paper showed that although exploitation of scene geometry information may augment semantic segmentation performance using Convolutional Neural Networks, it is not mandatory for geometry to be known or estimated during actual CNN deployment/inference. A novel regularizer was proposed that penalizes differences between semantic and depth predictions on presumed object boundaries during segmentor training. Neither ground-truth depth maps or special data (e.g., stereoscopic 3D) at the training stage, nor known or estimated depth maps at the inference stage are required. Quantitative evaluation was performed on a public scene parsing video dataset for autonomous driving. The results indicate that depth map utilization only during training, without resorting to a multitask setup which may demand a more complex backbone CNN to avoid underfitting, can be sufficient for increasing accuracy during deployment/inference. The obtained model predicts better semantic maps, with zero increase in computational requirements at the inference stage.

6. REFERENCES

[1] I. Mademlis, N. Nikolaidis, A. Tefas, I. Pitas, T. Wagner, and A. Messina, "Autonomous unmanned aerial vehicles filming in dynamic unstructured outdoor environments," *IEEE Signal Processing Magazine*, vol. 36, no. 1, 2018.

[2] S. Papadopoulos, I. Mademlis, and I. Pitas, "Neural vision-based semantic 3D world modeling," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2021.

[3] C. Symeonidis, E. Kakaletsis, I. Mademlis, N. Nikolaidis, A. Tefas, and I. Pitas, "Vision-based UAV safe landing exploiting lightweight Deep Neural Networks," in *Proceedings of the International Conference on Image and Graphics Processing (ICIGP)*. 2021, ACM.

[4] N. Yang, R. Wang, J. Stuckler, and D. Cremers, "Deep virtual stereo odometry: Leveraging deep depth prediction for monocular direct sparse odometry," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[5] Z. Liang, Y. Feng, Y. Guo, H. Liu, W. Chen, L. Qiao, L. Zhou, and J. Zhang, "Learning for disparity estimation through feature constancy," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[6] J. Bian, Z. Li, N. Wang, H. Zhan, C. Shen, M. Cheng, and I. Reid, "Unsupervised scale-consistent depth and ego-motion learning from monocular video," in *Proceedings of Advances in neural information processing systems (NIPS)*, 2019.

[7] M. Schwarz, A. Milan, A. S. Periyasamy, and S. Behnke, "Rgb-d object detection and semantic segmentation for autonomous manipulation in clutter," *The International Journal of Robotics Research*, vol. 37, no. 4-5, 2018.

[8] W. Wang and U. Neumann, "Depth-aware CNN for RGB-D segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[9] Y. Cao, C. Shen, and H. T. Shen, "Exploiting depth from single monocular images for object detection and semantic segmentation," *IEEE Transactions on Image Processing*, vol. 26, no. 2, 2017.

[10] J. Jiao, Y. Wei, Z. Jie, H. Shi, R.WH. Lau, and T.S. Huang, "Geometry-aware distillation for indoor semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[11] S. Zhu, G. Brazil, and X. Liu, "The edge of depth: Explicit constraints between segmentation and depth," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[12] P.Y. Chen, A. H Liu, Y.C. Liu, and Y.C.F Wang, "Towards scene understanding: Unsupervised monocular depth estimation with semantic-aware representation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[13] P. Z. Ramirez, M. Poggi, F. Tosi, S. Mattoccia, and L. Di Stefano, "Geometry meets semantics for semi-supervised monocular depth estimation," in *Proceedings of the Asian Conference on Computer Vision (ACCV)*. Springer, 2018.

[14] V. Guizilini, R. Hou, J. Li, R. Ambrus, and A. Gaidon, "Semantically-guided representation learning for self-supervised monocular depth," *arXiv preprint arXiv:2002.12319*, 2020.

[15] H. Jiang, G. Larsson, M. M. G. Shakhnarovich, and E. Learned-Miller, "Self-supervised relative depth learning for urban scene understanding," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[16] J. Novosel, P. Viswanath, and B. Arsenali, "Boosting semantic segmentation with multi-task self-supervised learning for autonomous driving applications," in *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, 2019.

[17] M. Klingner, A. Bar, and T. Fingscheidt, "Improved noise and attack robustness for semantic segmentation by using multi-task training with self-supervised depth estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020.

[18] L. Hoyer, D. Dai, Y. Chen, A. Köring, S. Saha, and L. Van Gool, "Three ways to improve semantic segmentation with self-supervised depth estimation," *arXiv preprint arXiv:2012.10782*, 2020.

[19] X. Huang, X. Cheng, Q. Geng, B. Cao, D. Zhou, P. Wang, Y. Lin, and R. Yang, "The apolloscape dataset for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018.

[20] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, "BiSeNet: Bilateral segmentation network for real-time semantic segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

INTELLIGENT ROAD SURFACE DEEP EMBEDDED CLASSIFIER FOR AN EFFICIENT PHYSIO-BASED CAR DRIVER ASSISTANCE

F. Rundo[1], R. Leotta[2], V. Piuri[3], A. Genovese[3], F. Scotti[3],S. Battiato[2]

[1]STMicroelectronics, ADG Central R&D
$francesco.rundo@st.com$
[2]University of Catania, IPLAB - Computer Science Department
$leotta.rob@gmail.com, battiato@dmi.unict.it$
[3]University of Milan, Computer Science Department
$\{vincenzo.piuri, angelo.genovese, fabio.scotti\}@unimi.it$

ABSTRACT

Car driving safety represents one of the major targets of the ADAS (Advanced Driver Assistance Systems) technologies deeply investigated by the scientific community and car makers. From intelligent suspension control systems to adaptive braking systems, the ADAS solutions allows to significantly improve both driving comfort and safety. The aim of this contribution is to propose a driving safety assessment system based on deep networks equipped with self-attention Criss-Cross mechanism to classify the driving road surface combined with a physio-based drowsiness monitoring of the driver. The retrieved driving safety assessment performance confirmed the effectiveness of the proposed pipeline.

Index Terms— ADAS, Automotive, Deep Learning, Road Classification, Intelligent Suspension

1. INTRODUCTION

The ADAS technologies are able to accomplish several tasks to assist the vehicle's driver leveraging different level of automation: from car driving assistance to fully autonomous driving or In-vehicle-Infotainment-Systems (IVIS) [1]. The recent ADAS technologies include automotive embedded systems suitable to provide ad-hoc warnings and alerts to the driver such as the Intelligent Speed Adaptation, collision warning systems or car driver drowsiness monitor [2]. Moreover, recent ADAS solutions combined visual information inside and outside the car with physiological assessment of the driver [2, 3]. In this context, the authors propose an innovative fully automated ADAS application which combines an efficient physiological car driver's drowsiness monitor driven by adaptive road surface risk assessment. The use of self-attention layers with temporal convolutional deep dilated architectures makes the proposed pipeline robust and efficient in monitoring driving risk. About road segmentation and classification, several approaches have been proposed. In [4] the authors described the development of a nice performer automated algorithms for extracting road features from Mobile Laser Scanning point cloud data. In [5] the authors proposed an interesting strategy to identify cracks on images captured during road pavement surveys. It adopted an efficient segmentation procedure, after appropriate image smoothing, followed by ad-hoc binary classification. Deep learning based solutions both supervised and unsupervised have been implemented for addressing the issue of a robust road segmentations [6, 7, 8, 9]. About driver attention monitoring systems, the authors have been deeply investigated the topic providing several scientific contributions and surveys [10, 11, 12, 13, 14, 15]. Several further approaches confirmed that physiological signal, especially the Photoplethysmography (PPG), can be efficiently used to monitor the car driver's attention level [15, 16, 17]. The proposed pipeline will be described in detail in the next paragraphs.

2. THE ROAD SURFACE SEGMENTATION AND CLASSIFICATION

The first sub-system of the proposed pipeline embeds a road segmentation and classification algorithm. In Fig. 1 is reported the scheme of the implemented approach. As schematized in Fig. 1, a Mask-R-CNN embedding a DenseNet-201 as feature generator backbone is proposed [18]. Mask-R-CNN is widely used in the automotive field [18]. The advantage of this architecture is that it provides a pixel-based segmentation of the driving frames as well as the generation of a bounding-box that characterizes the Region of Interest (ROI) on which to perform post-processing. The segmented road (ROI) will be fed as input of the enhanced downstream ResNet-101 in which we have embedded a Recurrent Criss-Cross Attention (RCCA) layer. The attention mechanism based on Criss-Cross processing was firstly proposed in [19] showing very promising performance in several tasks including semantic segmentation. More in detail, for each source image/feature

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

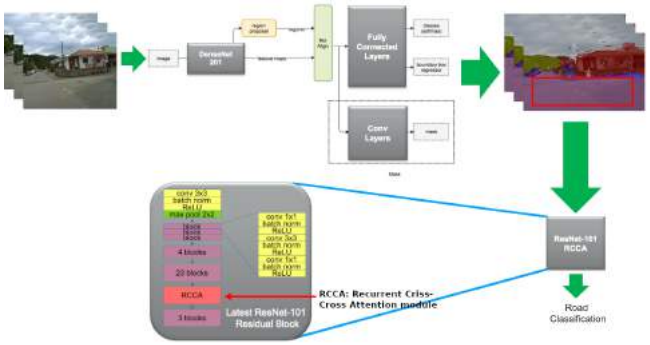


Fig. 1. The proposed Road Surface Classifier: Mask-R-CNN with a Recurrent Criss-Cross Attention (RCCA) enhanced ResNet-101

pixel, an innovative Criss-Cross attention module computes the contextual information of all the correlated pixels on its Criss-Cross path. This attention pre-processing combined with further recurrent operations allow the Criss-Cross method to leverage the full-image dependencies during the learning session of the deep network [19]. Lets formalize the attention processing embedded in the Criss-Cross layer we have implemented. Given a local feature map $H \in R^{C\times W\times H}$ where C is the original number of channels while $W \times H$ represents the spatial size of the generated feature map trough a Deep Convolutional Network. The Criss-Cross layer applies two preliminary 1×1 convolutional layers on H in order to generate two feature maps F_1 and F_2, which belong to $R^{C'\times W\times H}$ and in which C' represents the reduced number of channels due to dimension reduction with respect to original (C). Lets define an *Affinity* function suitable to generate the Attention-Map $A_M \in R^{(H+W-1)\times(W\times H)}$. The Affinity operation is so defined. For each position u in the spatial dimension of F_1, we extract a vector $F_{1,u} \in R^{C'}$. Similarly, we define the set $\Omega_u \in R^{(H+W-1)\times C'}$ by extracting feature vectors from F_2 at the same position u. So that, $\Omega_{i,u} \in R^{C'}$ is the i-th element of Ω_u. Taking into account the above operations, we can define the introduced *Affinity* operation as follows:

$$\delta^A_{i,u} = F_{1,u}\Omega^T_{i,u} \tag{1}$$

where $\delta^A_{i,u} \in D$ is the affinity potential i.e. the degree of correlation between features $F_{1,u}$ and $\Omega_{i,u}$ for each $i = [1, ..., H + W - 1]$, and $D \in R^{(H+W-1)\times(W\times H)}$. Then, we apply a softmax layer on D over the channel dimension to calculate the attention map A_M. Finally, another convolutional layer with 1×1 kernel will be applied on the feature map H to generate the re-mapped feature $\vartheta \in R^{C\times W\times H}$ to be used for spatial adaptation. At each position u in the spatial dimension of ϑ, we can define a vector $\vartheta_u \in R^C$ and a set $\Phi_u \in R^{(H+W-1)\times C}$. The set Φ_u is a collection of feature vectors in ϑ having the same row or column with position u. At the end, the final contextual information will be obtained by an *Aggregation* operation defined as follows:

$$H'_u = \sum_{i=0}^{H+W-1} A^{i,u}_M \Phi_{i,u} + H_u \tag{2}$$

where H'_u is a feature vector in $H' \in R^{C\times W\times H}$ at position u while $A^{i,u}_M$ is a scalar value at channel i and position u in A_M. The so defined contextual information H'_u is then added to the given local feature H to augment the pixel-wise representation and aggregating context information according to the spatial attention map A_M. The Criss-Cross processing fails to process the connections there are among one pixel and its around. For this reason, a Recurrent Criss-Cross processing was embedded in the proposed pipeline (with $R = 2$ i.e. Criss-Cross operations can be unrolled into 2 loops)[19]. In order to enhance the deep classifier, we have included a Criss-Cross layer in the latest residual block of the ResNet-101 as reported in Fig. 1. The proposed pipeline has been trained and tested in the RTK dataset [20] trying to discriminate four types of road: asphalt, paved, potholes and unpaved. The output of the Criss-Cross enhanced ResNet-101 (having a softmax as latest layer) is a binary mask of four bits [asphalt, paved, potholes, unpaved]. The bits set to 1 confirm that the segmented input frames contains this kind of road surface. The performance results will be showed in the next sections.

3. THE PPG BASED CAR DRIVER DROWSINESS MONITORING SYSTEM

As introduced, the second block of the proposed ADAS pipeline is the physio-based car driver drowsiness monitoring system. Specifically, we proposed a car-driver attention level monitoring based on the usage of the driver's Photoplethysmographic (PPG) signal. Through a deep PPG signal features analysis [3] a non-invasive blood volume dynamic assessment can be retrieved. More in detail, a common PPG waveform consist of a pulsatile physiological signal ('AC') that embedding blood volume information overlapped with slowly varying component ('DC') that represents information correlated to the skin tissues (where the PPG sensor is placed), respiration and thermoregulation. With a device consisting in a light-emitter and a detector placed on the skin that measure the amount of light either transmitted or reflected we can detect the blood volume changes occuring with the heart pressure pulse. The correlation between blood volume changes and the Autonomic Nervous System that manage the alert levels of the subject and cardiac activity allows to consider the PPG an excellent indirect detector of the subject's level of attention [3, 10, 11]. In addition, the correct level of attention required for safe driving is computed and adjusted according to the driving context (speed, pavement conditions, adjacent vehicles, and so on) [14]. In this work, the Silicon Photomultiplier sensor [10, 11] was used as PPG detector.

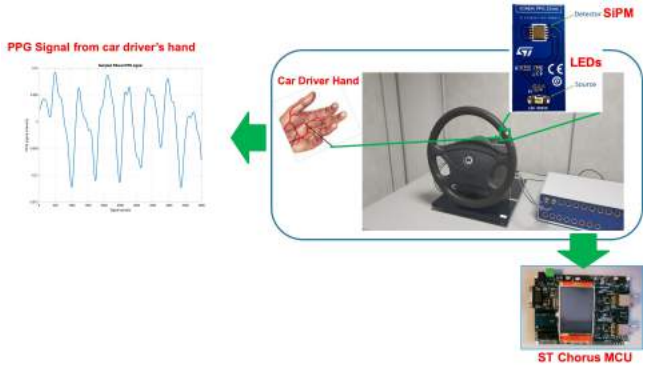


Fig. 2. The PPG sensing device embedded in the car steering.

More in detail, the suggested PPG probes consists in a large area of n-on-p Silicon Photomultipliers (SiPMs) fabricated at STMicroelectronics [10, 11]. The SiPMs array has a total area of $4.0 \times 4.5mm^2$ and 4871 square microcells with $60\,\mu\mathrm{m}$ pitch, packaged in a surface mount housing (SMD) with about $5.1 \times 5.1mm^2$ total area [11]. Furthermore, on the SMD package was glued a Pixelteq dichroic bandpass filter by means the use of Loctite 352TM adhesive. The aforementioned bandpass filter was set with a pass band centered at about $540\,\mathrm{nm}$ with a Full Width at Half Maximum (FWHM) of $70\,\mathrm{nm}$ and an optical transmission higher than $90 - 95\%$ in the pass band range. As light emitter we have used the OSRAM LT M673 LEDs in SMD package that emits at $830\,\mathrm{nm}$ which is based on InGaN technology [11]. More in detail, the aforementioned LEDs devices have an area of $2.3 \times 1.5mm^2$, spectral bandwidth of $33\,\mathrm{nm}$, viewing angle of $120°$ and lower power emission (mW) in the standard operation range. The authors have designed a printed circuit board (PCB) in order to make the PPG probe easily to use. More implementation details can be found in [11]. In Fig. 2 we report an overall scheme of the proposed PPG sensing framework (SiPM + LEDs) embedded in the car steering. As reported in Fig. 2, the filtering and stabilization of the raw PPG signal collected from the car driver hand will be performed by the developed algorithms running as firmware in the ST Chorus MCU [11, 12, 13, 14, 15, 16]. After that, the hyper-filtering approach we have implemented and patented [11], [16] will be applied to the collected stabilized PPG data in order to retrieve an evaluation of the attention level of the driver from which the PPG signal is sampled. We have configured the hyper-filtering approach for the application herein described. Specifically, the proposed hyper-filtering system has been inspired by the widely-accepted idea of hyper-spectral processing used in 2D imaging [16]. Hyper-spectral imaging gather visual information through the whole electromagnetic spectrum, in order to retrieve the so called "frequency spectrum of each pixel" [15]. Thus, using the same method, the authors considered the information set retrieved from such "hyper-filtered" signals i.e. the set of signals obtained by different frequency filtering of the source time-serie (PPG in our use-case). With the proposed hyper-filtering approach we are able to collect valuable information about the frequency spectrum of the car driver's PPG signal and then about the correlated driver's attention level (Drowsiness monitoring). More in detail, we have divided the valuable PPG frequency range $0.5\,\mathrm{Hz}$–$10\,\mathrm{Hz}$ in several sub-ranges in which we have applied the Butterworth pass-band filter (high-pass and low-pass filters) as described in [16, 15]. Thus, we have configured two layers of hyper-filtering systems which are able to modulate the frequencies in the low-pass application, meanwhile preserving the cut-off frequency of the high-pass filter (Hyper low-pass filtering layer) and vice-versa (Hyper high-pass filtering layer). The applied hyper-filtering frequency setup is reported in Table 1 and Table 2. We proposed the usage of Butterworth filters in both hyper-filtering setup since they do not create modulations or distortions in the bandwidth [14, 15, 16]. We retrieved the frequency values reported in Table 1 and Table 2 through a Reinforcement Learning algorithm with a reward function correlated to the car driver drowsiness classification accuracy [15, 16]. Once

F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
HP	0.5	/	/	/	/	/	/	/	/	/	/
LP	0.0	1.1	3.2	3.5	3.8	3.9	4.0	4.1	5.0	5.1	6.3

Table 1. Hyper Low-pass filtering setup (in Hz).

F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
HP	0.1	1.1	2.3	2.4	3.2	3.5	4	4.2	5	5.3	6.4
LP	7.0	/	/	/	/	/	/	/	/	/	/

Table 2. Hyper High-pass filtering setup (in Hz).

the hyper-filtering setup has been assessed, the collected car driver PPG raw signal will be processed accordingly. Specifically, from the collected source PPG driver signal, a subset of hyper-filtered signals will be generated through the frequency setup as per Table 1 and Table 2. Lets define $W_i^{PPG}(t,k)$ the single segmented waveform of the i-th hyper-filtered PPG time-serie. For each sample $s^i(t_k)$ of the segmented PPG waveform $W_i^{PPG}(t,k)$, we will compute a signal-pattern representing the dynamic of the sample $s^i(t_k)$ for each i-th $W_i^{PPG}(t,k)$ waveforms. Consequently, we collect a large dataset of hyper-filtered signal patterns [14, 15, 16]. As soon as the driver put the hand over the PPG sensing probe embedded on the steering wheel, the hyper-filtering pipeline starts to work generating the signal-patterns to be fed as input to the Deep Learning block as detailed in Fig. 3. Specifically, the designed classifier is a Deep 1D Temporal Dilated Convolutional Neural Network (1D-TCNN) with residual blocks [15]. The temporal convolutional network is mainly characterized by a causal convolution layer [15]. The designed 1D-TCNN is composed as follows: 25 residual blocks with 3×3 kernel filters, where such of them contains dilated

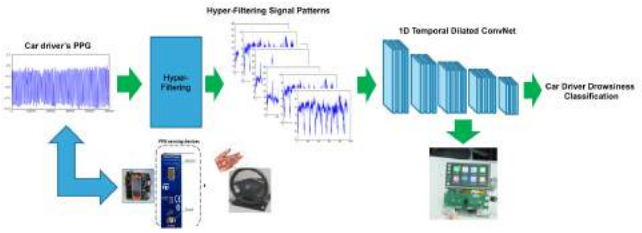


Fig. 3. The Physio-based Car Driver Drowsiness Monitoring System

convolution operations in which the dilation factor size starts from 2 and increase (power of 2) till to 16, normalization, ReLU activation, spatial dropout layers and a downstream softmax layer. The so designed 1D-TCNN is able to classify the input hyper-filtered PPG patterns coming from a drowsy or wakeful driver ($(0.0 - 0.5)$, $(0.51 - 1.0)$ respectively). As reported in Fig. 3, the designed 1D TCNN is running over the STA1295A Accordo5 MCU [15, 16].

4. EXPERIMENTAL RESULTS AND CONCLUSION

We tested the implemented pipeline, validating the single subsystems and then arranging a composite scenario of a road surface-driven risk assessment (driving safety monitoring system). Specifically, we considered the following risk assessment: asphalt (low driving risk), paved (medium driving risk due to certain braking problems) unpaved / potholes (high driving risk). More in detail, if high or medium risk level is assessed by the road classification system (Mask-R-CNN with Criss-Cross ResNet-101 downstream) the driving safety monitoring system will check that 1D-TCNN confirms a corresponding "wakeful" attention classification. Otherwise, acoustic alert-signal will be emitted by the Audio underlying System (STA1295A MCU) in order to alert the drowsy driver. If the driver's PPG signal is not available for some reasons (for instance: the driver does not have his/her hand placed over the PPG sensing devices in the steering wheel), the authors have developed a system for ad-hoc visual reconstruction of the PPG signal by means of an innovative motion magnification technique applied to specific driver's facial landmarks [12]. Now, follows more details about the performance of the proposed sub-systems. About the proposed road surface classification deep pipeline, we validated and compared our pipeline using the RTK dataset and related algorithms [20, 9]. We arranged the dataset into 80% for training and validation while the remaining 20% for testing. The following Table 3 shows some performance benchmarks. Regarding the driver's physio-based drowsiness assessment, we have tested the suggested pipeline by gathering various PPG measurements of several subjects in different scenarios (Drowsy vs Wakeful drivers) under clinical study covered by the *Ethical Committee CT1 authorization.113/2018/PO*.

Method	**Road Surface Classification Performance**		
	Low Risk (Asphalt)	**Medium Risk (Paved)**	**High risk (unpaved / potholes)**
Proposed	**93%**	92%	**97% / 97%**
Proposed w/o Criss-Cross	92%	89%	89% / 92%
Proposed w/o ResNet-101	88%	88%	84% / 82%
[20, 9]	92%	**94%**	94% / **97%**

Table 3. Road Surface Classification Performance.

We collected data from 70 patients with different features such as ages, gender, etc. Furthermore, simultaneously with the PPG signals we also acquired EEG signal to be able to verify the attention level (alpha and beta waves) [13]. We have sampled the PPG signals of the subjects by means of the hardware setup detailed in this contribution with a sampling frequency equal to $1\,\mathrm{kHz}$. We gathered 5 minutes of PPG signals for both condition (Drowsiness and Wakefulness). The so collected PPG time-series, have been organized as follow: 70% was used for the training and validation phases while the remaining 30% was used for testing. For the training phase of the 1D-TCNN we set an initial learning rate equal to 0.001 and dropout factor equal to 0.5. Furthermore, a classic SGD algorithm was used. The following Table 4 reports the performance obtained with the aforementioned pipeline compared to similar pipeline based on deep learning [16]. The collected performance results (related to the single

Method	**Driver Drowsiness Monitoring**	
	Drowsy Driver	**Wakeful Driver**
Proposed	**98.71%**	**99.03%**
[16]	96.50%	98.40%

Table 4. Car Driver Drowsiness Classification Performance.

subsystems) confirm that the overall proposed pipeline performs very well allowing an adaptive, robust and innovative fully automated assessment of the driving risk based on road surface classification. As confirmed by the results reported in Table 3, the use of Criss-Cross enhanced downstream ResNet-101 classifier allow to obtain significantly improvement in terms of classification performance. This research was supported by the National Funded Program 2014-2020 under grant agreement n. 1733, (ADAS + Project).

5. REFERENCES

[1] Ryosuke Okuda, Yuki Kajiwara, and Kazuaki Terashima, "A survey of technical trend of adas and autonomous driving," in *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test*. IEEE, 2014, pp. 1–4.

[2] Chang Wang, Qinyu Sun, Yingshi Guo, Rui Fu, and

Wei Yuan, "Improving the user acceptability of advanced driver assistance systems based on different driving styles: A case study of lane change warning systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4196–4208, 2019.

[3] Nicoleta Minoiu Enache, Saïd Mammar, Mariana Netto, and Benoit Lusetti, "Driver steering assistance for lane-departure avoidance based on hybrid automata and composite lyapunov function," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 1, pp. 28–39, 2009.

[4] Haiyan Guan, Jonathan Li, Yongtao Yu, Michael Chapman, and Cheng Wang, "Automated road information extraction from mobile laser scanning data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 194–205, 2014.

[5] Henrique Oliveira and Paulo Lobato Correia, "Road surface crack detection: improved segmentation with pixel-based refinement," in *2017 25th IEEE EUSIPCO Proceedings*. IEEE, 2017, pp. 2026–2030.

[6] Christian Koch, Kristina Georgieva, Varun Kasireddy, Burcu Akinci, and Paul Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 196–210, 2015.

[7] Jin Tian, Jiazheng Yuan, and Hongzhe Liu, "Road marking detection based on mask r-cnn instance segmentation model," in *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*. IEEE, 2020, pp. 246–249.

[8] Shashank Yadav, Suvam Patra, Chetan Arora, and Subhashis Banerjee, "Deep cnn with color lines model for unmarked road segmentation," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 585–589.

[9] Thiago Rateke and Aldo von Wangenheim, "Road surface detection and differentiation considering surface damages," *Autonomous Robots*, pp. 1–14, 2021.

[10] Vincenzo Vinciguerra, Emilio Ambra, et al., "Ppg/ecg multisite combo system based on sipm technology," in *Convegno Nazionale Sensori*. Springer, 2018, pp. 353–360.

[11] Francesco Rundo, Sabrina Conoci, Alessandro Ortis, and Sebastiano Battiato, "An advanced bio-inspired photoplethysmography (ppg) and ecg pattern recognition system for medical assessment," *Sensors*, vol. 18, no. 2, pp. 405, 2018.

[12] Francesca Trenta, Sabrina Conoci, Francesco Rundo, and Sebastiano Battiato, "Advanced motion-tracking system with multi-layers deep learning framework for innovative car-driver drowsiness monitoring," in *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*. IEEE, 2019, pp. 1–5.

[13] Francesco Rundo, Sergio Rinella, et al., "An innovative deep learning algorithm for drowsiness detection from eeg signal," *Computation*, vol. 7, no. 1, pp. 13, 2019.

[14] Francesco Rundo, Sabrina Conoci, Sebastiano Battiato, et al., "Innovative saliency based deep driving scene understanding system for automatic safety assessment in next-generation cars," in *2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive*. IEEE, 2020, pp. 1–6.

[15] Francesco Rundo et al., "Advanced 1d temporal deep dilated convolutional embedded perceptual system for fast car-driver drowsiness monitoring," in *2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive*. IEEE, 2020, pp. 1–6.

[16] Francesco Rundo, Concetto Spampinato, and Sabrina Conoci, "Ad-hoc shallow neural network to learn hyper filtered photoplethysmographic (ppg) signal for efficient car-driver drowsiness monitoring," *Electronics*, vol. 8, no. 8, pp. 890, 2019.

[17] Hyeonjeong Lee, Jaewon Lee, and Miyoung Shin, "Using wearable ecg/ppg sensors for driver drowsiness detection based on distinguishable pattern of recurrence plots," *Electronics*, vol. 8, no. 2, pp. 192, 2019.

[18] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[19] Zilong Huang, Xinggang Wang, Lichao Huang, et al., "Ccnet: Criss-cross attention for semantic segmentation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 603–612.

[20] Thiago Rateke, Karla Aparecida Justen, and Aldo von Wangenheim, "Road surface classification with images captured from low-cost camera-road traversing knowledge (rtk) dataset," *Revista de Informática Teórica e Aplicada*, vol. 26, no. 3, pp. 50–64, 2019.

AUTOMATED PARKING TEST USING ISAR IMAGES FROM AUTOMOTIVE RADAR

Neeraj Pandey, Shobha Sundar Ram

Indraprastha Institute of Information Technology Delhi, New Delhi 110020 India
E-mail: {neerajp, shobha}@iiitd.ac.in

ABSTRACT

Automated driving tests using cameras have been researched for expediting the training and testing of car drivers. We propose an automated parking test using millimeter-wave automotive radars. The advantage is that these radars can be operated even in low visibility conditions. We propose generating high-resolution inverse synthetic aperture radar (ISAR) images of a vehicle under test (VUT) parking into a designated parking slot from an externally mounted radar. The trajectory of the motion is estimated from the ISAR data using polynomial curve fitting from which the VUT is deemed to have either correctly or incorrectly parked. We experimentally validate the proposed method with millimeter-wave radar data gathered for cars performing perpendicular and 45° angle parking.

Index Terms— ISAR, Parking test, Automotive radar

1. INTRODUCTION

A key component towards improving road safety is to have well-trained and tested automobile drivers on the road. Recently, there has been research and development of automated driving license tests in order to address the costs, man-hours and other types of inadequacies associated with manual driving tests such as the tedium and corruption of the driving test inspectors. Driving tests typically involve testing the drivers in several key components such as lane changing, turning, following road rules, and parking. Currently, cameras have been the sensors of choice for automotive license tests [1]. In this paper, we propose the use of millimeter-wave radars specifically for conducting automated parking tests. The main advantage of radar over other sensors such as cameras is that they can be used in low light/visibility conditions. Frequency modulated continuous wave radars are currently heavily used in automotive systems for pedestrian detection, blind spot detecting, object classification, and collision avoidance [2]. In our work, we propose that we use high-resolution radar images to test if the driver is following the designated test trajectory to the final parking spot.

Concurrent to automatic license testing development are the parking assistance systems on automobiles using cameras, lidars, and radars. These systems have been used to detect occupancy of parking slots [3, 4] and in aiding the drivers in the correct positioning of the vehicle, especially while backing in [5]. The main distinction between the system that we propose to the existing parking assistance systems is that the radar is mounted *outside* the vehicle. Thus the objective is to *test/train* the driver as opposed to *assisting* the driver to complete the parking. Second, conventional automotive radars provide point cloud information of the target scatterers in the radar field-of-view. Each scatterer is marked with associated features of range, Doppler velocity, and signal strength. High-end systems consisting of multiple receiver channels provide additional azimuth and elevation information as well. In the proposed system, instead of generating point cloud data of the scatterers on the vehicle, we consider a low cost and complexity, single-channel wideband radar, and generates high-resolution range-crossrange images using inverse synthetic aperture radar (ISAR) imaging techniques.

ISAR imaging is an established technique for generating high-resolution radar images of turning targets using single-channel radar data and has been extensively researched for airborne, waterborne, and ground-based vehicles [6]. ISAR, essentially, exploits the turning motion of a target to synthesize a large cross-range aperture in order to obtain fine cross-range resolution. In the past few years, there have been significant studies of both synthetic aperture radar (SAR) [7], and ISAR imaging of automotive targets using wideband data [8, 9, 10] for target detection and classification. These works have shown that these types of radar images provide rich information of the size, orientation, and trajectory of dynamic targets [10, 11]. The main challenges of ISAR are twofold: First, the translational motion parameters of the automotive targets must be properly estimated in order to calibrate/compensate for their translational motion; second, the angular turning velocity must be accurately measured in order to obtain accurate cross-range estimates. In the parking test that we propose, we specifically address these challenges by determining if the parking is correct based on the estimated two-dimensional trajectory followed by the vehicle under test (VUT) using the ISAR images. We have experimentally validated our proposed parking test using measurement radar data gathered from the 77 GHz Texas Instruments AWR-1843 automotive radar. Wide-band single channel radar data gathered along the target trajectory are suitably processed to obtain ISAR images from which the trajectory of the target is estimated.

We can summarize our main contributions in this paper as follows- First, we have proposed using ISAR images generated from a single channel externally mounted automotive radar for two types of parking - angle parking and perpendicular parking; Second, we have developed an automated parking test algorithm that uses these ISAR images to test if the parking maneuver is correct. Our paper is organized as follows. In the following section, we present the theory of radar signal, signal processing, and the parking test algorithm. In Section.3, we present the details of the experimental set up followed by the results in Section 4. Finally, we conclude the paper in Section 5.

2. THEORY

Radar Signal: We assume that the VUT is moving on the XY ground plane with the height along the Z axis. The automotive radar is mounted outside of the car in a fixed position at the origin. The radar is operated from a short-range from the VUT in a line-of-sight environment. The radar transmits a frequency modulated continuous wave radar signal as given by

$$S_{tx}(\tau) = rect\left(\frac{\tau}{T_{PRI}}\right) e^{j2\pi f_c \tau} e^{j\pi K \tau^2}, \quad (1)$$

where f_c is the carrier frequency, and the K is the chirp rate of the transmitted signal. In (1), $rect\,(\cdot)$ indicates that the transmitting signal is defined for the pulse repetition interval T_{PRI}. At short ranges, we consider the VUT as an extended target with B point scatterers. We assume that the amplitude of each point scatterer, a_b, slowly fluctuates and hence is constant within a coherent processing interval (T_{CPI}). Each point scatterer is dynamic, and the time-varying range of the scatterer is $r_b(t) = R_b + v_b t$, where R_b is the starting distance from the radar and v_b is the relative radial velocity with respect to radar. Here, t is the slow time across multiple pulses comprising one T_{CPI}. Due to the motion of the scatterer, the backscattered radar signal is Doppler shifted by $f_{D_b} = \frac{2v_b f_c}{c}$, where c is the velocity of light. The down-converted radar received signal from the target can be expressed in terms of fast (τ) and slow time as

$$S_{rx}(\tau, t) = \sum_{b}^{B} a_b(t) rect\left(\frac{\tau - \frac{2r_b}{c}}{T_{PRI}}\right) e^{j2\pi f_c \frac{2r_b(t)}{c}} e^{j\pi K(\tau - \frac{2r_b(t)}{c})^2} + \mu, \quad (2)$$

where μ is the additive receiver noise. The radar signal is sampled at sampling frequency f_s resulting in N fast time samples in every PRI and M slow time samples in every T_{CPI}. Thus, the discrete form of (2) is given by

$$S_{rx}[n, m] = \sum_{b}^{B} a_b[m] rect\left[\frac{n - n_b}{N}\right] e^{-j\frac{4\pi f_c}{c} R_b} e^{-j2\pi m f_{D_b} T_{PRI}} e^{j\pi K \frac{1}{f_s{}^2}(n - n_b)^2} + \mu, \quad (3)$$

where n_b is integer rounded from $\frac{2r_b(t)}{cf_s}$. The Doppler shift in (3) is due to both the translational as well as rotation motion of the target. Our first step in processing the digitized radar data is to perform translational motion compensation within each T_{CPI} based on [12, 6]. In the interests of space, we are not discussing these steps in complete detail. The range compensated signal is processed using the two-dimensional (2D) Fourier transform along the fast and slow time dimensions to obtain the range-Doppler ambiguity plot, as shown in

$$\chi[r, f_d] = 2DFT\{S_{rx}[n, m]\}. \quad (4)$$

We map the Doppler axis in the ambiguity plot to the crossrange - axis using an estimate of the angular velocity ω of the target for the corresponding T_{CPI} as shown in

$$\chi[r, f_d] \overset{f_d \times \frac{\lambda_c}{2\omega}}{\rightarrow} \chi[r, cr]. \quad (5)$$

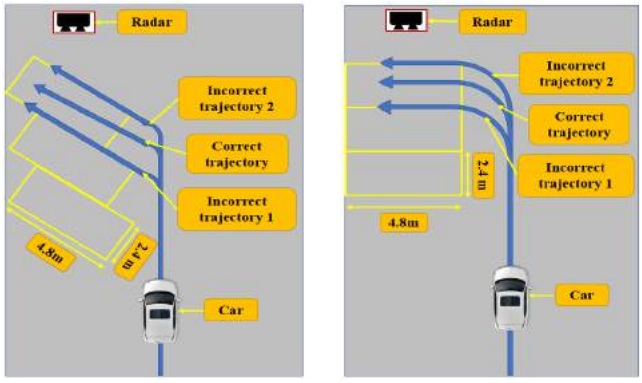
The assumption is that the angular velocity is constant during the T_{CPI}. The Doppler axis f_d is multiplied by $\frac{\lambda_c}{2\omega}$ where λ_c is the wavelength corresponding to f_c in free space. There are many ways to estimate the angular velocity, ω, of a cooperative target such as the use of additional sensors such as gyrometers and accelerometers. In this work, we estimate the ω for every T_{CPI} from the change in the yaw, θ, over consecutive intervals, where $\theta[m]$ is

$$\theta[m] = \arctan\left(\frac{y_C[m] - y_C[m-1]}{x_C[m] - x_C[m-1]}\right). \quad (6)$$

Here (x_C, y_C) are the coordinates of the centroid of the VUT chassis with respect to the radar. We compute these coordinates based on the initial location of the VUT, the predefined trajectory, and the duration that the VUT takes to complete the trajectory.

Parking Test Algorithm: In the parking test, we infer whether the VUT has been appropriately parked based on the estimated trajectory of the VUT. In the *training algorithm*, we perform the following steps: First, we collect L ISAR images of the car following a correct trajectory into the designated parking slot. Then we collect a similar set of ISAR images of the car following several incorrect trajectories resulting in parking outside of the designated parking slot. Then for each trajectory (both correct and incorrect), we convert the images to gray-scale. Then we choose a 2D bounding box comparable to the length and width of the car and slide the center of the box across the pixels of each l^{th} image while keeping the dimensions of the box fixed. For each sliding position of the bounding box, we compute the sum of the energy in the corresponding pixels bounded by the box.Then, we identify the 2D pixel position, x_c^l, y_c^l which has the maximum energy as the dominant scattering center position for that l^{th} image. We repeat these steps across all L ISAR images corresponding to the VUT motion. Then we curve fit a 2D polynomial across the dominant scatterer position to estimate the trajectory of the VUT across all the L images. We hypothesize that this two-dimensional polynomial function will correspond approximately to the trajectory of the center of the car. We repeat the exercise for the motion of the car along the incorrect trajectories. These polynomial functions are stored and used while testing.

During the *test*, similar polynomials are estimated for the VUT motion and compared to the polynomials generated from the training. Then the VUT parking motion is deemed to be either correct or incorrect based on the closest fit of the test polynomial to the training polynomials. The advantages of this simple parking test are that the speed of VUT need not be identical to those of the car used for training. Hence, auxiliary sensors for estimating the translational motion characteristics of the test vehicles are not required. Further, the size of the VUT can differ a little from that used during training by adjusting the size of the bounding box.



(a) (b)

Fig. 1. Three trajectories for the parking test for (a) 45° angle parking and (b) perpendicular parking.

3. EXPERIMENTAL SET UP

We use the Texas Instruments AWR 1843, a 77GHz millimeter-wave radar, for experimental data collection. We configure the radar to transmit a frequency modulated continuous wave signal with the parameters listed in Table 1 to operate as a short-range radar.The transmitted radar signal's pulse repetition interval (T_{PRI}) is set to be 400

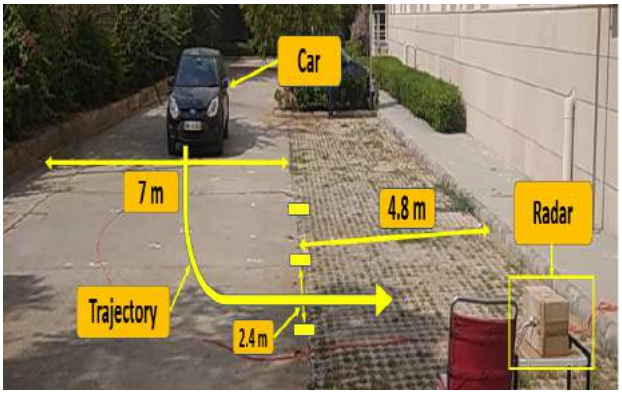


Fig. 2. The experiment setup for the parking test radar is situated at the location (0,0,0.5)m, and the car is parking at perpendicular parking.

Table 1. Automotive radar TI-AWR 1843 parameters for generating ISAR images

Parameters	Values
Carrier frequency (f_c)	77GHz
Sampling Frequency (f_s)	5MHz
Bandwidth (BW)	2GHz
Chirp rate (K)	7.5×10^{12} Hz^2
Chirp duration (T_{PRI})	400μs
Coherent processing interval (T_{CPI})	0.1s
Ramp time	267 μs
Idle time	133 μs
Transmitted power (P_t)	14dBm

μs, with a 133 μs idle time and a 267 μs ramp time. The chirp factor, or slope, is set to 7.532 MHz/μs, giving it a bandwidth of 2GHz, resulting in 7.5 cm range resolution. The sampling frequency is selected as 5MHz, resulting in 1328 fast time samples in each T_{PRI}. A single coherent processing interval, T_{CPI}, of 0.1s duration is formed from slow time data of 250 chirps.

We have performed our experiments with two commonly used parking scenarios - the 45° angle parking and perpendicular parking as shown in Fig.1. The experimental setup for the parking test is shown in Fig.2, in which the millimeter-wave radar is situated at the origin. The dimensions of the trajectory and parking lot for conducting the parking test are selected as per the standard defined by government agencies for parking [13]. The road along the parking trajectory is 7m in width. In the angle parking case, the car must first take a straight path, and then it should be parked at 45° from the road in a parking slot that is 2.4×4.8 meters. In perpendicular parking, the car must take a straight path and then parked at 90° from the road into a parking slot that is also of the same dimensions as the previous case. For our experimental data collection, we have chosen three trajectories for each type of parking test - one correct parking trajectory and two incorrect parking trajectories. We use two small-size cars for our experiment, The first car is a Ford Figo of $3.9 \times 1.7 \times 2.5$ meters size, and the second car is a Honda Brio of a comparable size of $3.6 \times 1.7 \times 1.5$ meters.

4. RESULTS

In this section, we will first discuss the ISAR images obtained while a car is being parked in the designated parking slot. We captured measurement data for 5 seconds for each parking test. We generate 50 ISAR images for each parking test using a T_{CPI} of 0.1 seconds. In Fig 3a, we show the ISAR images for the perpendicular parking test for the Ford Figo. In each row, four images are shown corresponding to T_{CPI} equal to 3.5, 4.0, 4.5, and 5.0 s. The images from the straight-line motion of the car are not shown in these results. The first row corresponds to the ISAR images generated when the car is taking the correct trajectory and parked correctly. The second and third rows show the ISAR images for the case when the car is taking the two incorrect trajectories shown in Fig1(b) and (c); from these ISAR images, we observe the car's dimensions along with the range and cross-range. Also, in the first few view-graphs (i,ii,v and vi), we observe that the car is oriented such that the longer dimension is along with the range while the shorter is along the cross-range. Then the car undergoes a turn such that the longer dimension is along the cross-range. Along the cross-range, we also observe the micro-Doppler tracks due to the wheels of the car. In all the figures, the cross-range axis varies because this axis depends on the target's rotational velocity ω, which changes in every T_{CPI}. Our hypothesis is that based on the intensity of the ISAR image pixels corresponding to the car, we will be able to determine if the car followed the designated trajectory into the correct parking slot. Next, we show the results for the angle parking for the Ford Figo in Fig3b for the same four-time instants. Again, we show the results for the correct parking (top row), and incorrect parking due to motion along the wrong trajectories (middle and bottom rows). Then, we repeated the measurements for the Honda Brio, and show the ISAR images for perpendicular parking in Fig.4a and angle parking in Fig.4b. The top rows for both figures correspond to the case when the car executed the correct parking, while the remaining two rows show the ISAR Images when the car performs incorrect parking. Just as in the previous case, the ISAR images give information about the car's size, position and its orientation along the range and cross-range axes. In Table 2 we report the results of the parking test.We use the Ford Figo data for training and Honda Brio data for tests. The first three trajectories are for the perpendicular parking case, while the remaining three are for the angle parking. Trajectories 1 and 4 correspond to the correct parking case, while the remaining correspond to incorrect parking. Table 2 shows the normalized mean square error (NMSE) between the estimated trajectory of the test case with each of the three training cases. The results indicate that the predicted trajectory from the algorithm is matched correctly to the ground truth in all cases based on the minimum NMSE. Thus the VUT has been correctly deemed to have either passed the parking test (cases 1 and 4) or failed the test (cases 2, 3, 5 and 6).

Table 2. Result of parking test algorithm using Ford Figo data for training and Honda Brio data for test.

	Predicted Trajectory					
True Trajectory	1	2	3	4	5	6
1	**0.332**	0.387	0.347	-	-	-
2	0.361	**0.347**	0.365	-	-	-
3	0.467	0.481	**0.466**	-	-	-
4	-	-	-	**0.512**	0.534	0.615
5	-	-	-	0.558	**0.512**	0.528
6	-	-	-	0.509	0.540	**0.411**

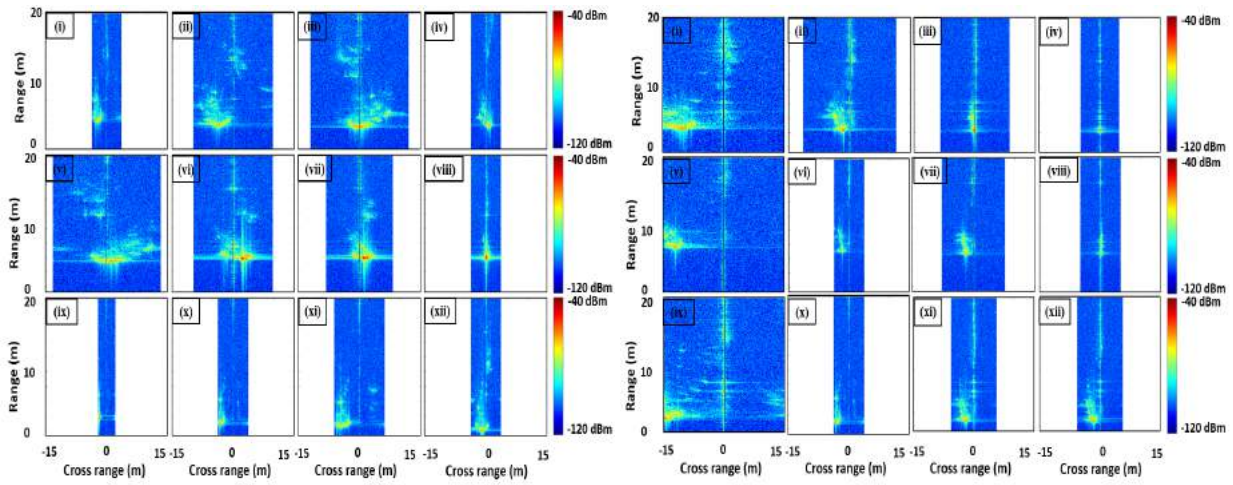


Fig. 3. ISAR images of Ford Figo carrying out (a) perpendicular parking and (b) angle parking, at 3.5, 4.0, 4.5, 5s. (1-iv) Top, (v-viii) middle, and (ix-xii) bottom rows in both images are generated for car following correct trajectory, incorrect trajectory-1, and incorrect trajectory-2 respectively.

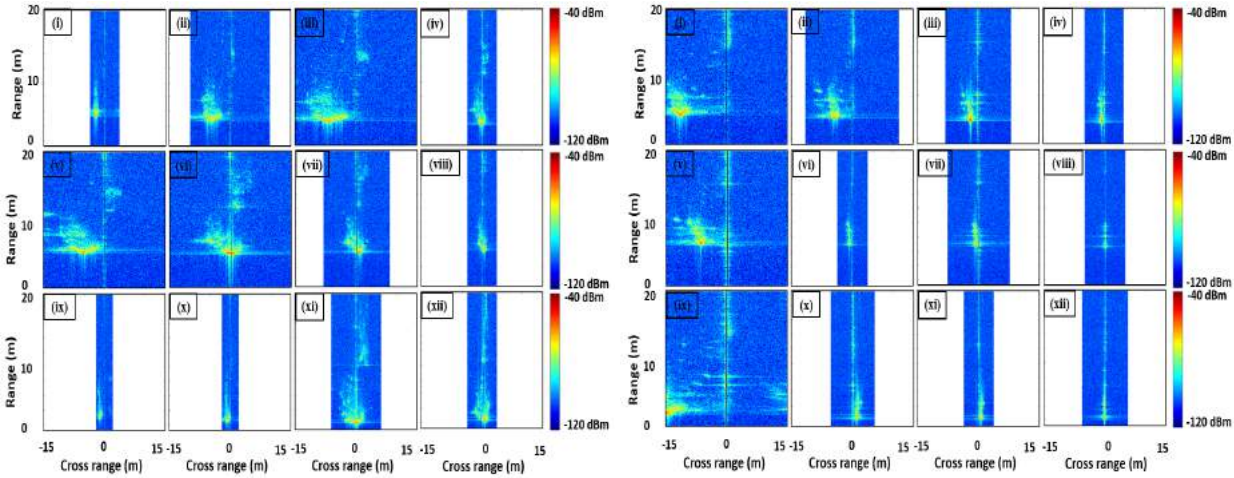


Fig. 4. Bottom row shows ISAR images of Honda Brio carrying out (a) perpendicular parking and (b) angle parking, at 3.5, 4.0, 4.5, 5s. (i-iv) Top, (v-viii) middle, and (ix-xii) bottom rows in both images are generated for car following correct trajectory, incorrect trajectory-1, and incorrect trajectory-2 respectively.

5. CONCLUSION

We have demonstrated the use of an externally mounted automotive radar to conduct automatic parking tests based on ISAR radar images. We have developed an automated parking test algorithm to classify the parking maneuvers of a VUT as either correct or incorrect based on the estimated trajectory of the VUT from the ISAR images. We have experimentally validated our proposed parking test algorithm using the data gathered using the TI-AWR 1843 millimeter-wave sensor.

6. REFERENCES

[1] A. U. Nambi, I. Mehta, A. Ghosh, V. Lingam, and V. N. Padmanabhan, "Alt: towards automating driver license testing using smartphones," in *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*, 2019, pp. 29–42.

[2] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 ghz frequency band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 845–860, 2012.

[3] T. Lin, H. Rivano, and F. Le Mouël, "A survey of smart parking solutions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3229–3253, 2017.

[4] V. Paidi, H. Fleyeh, J. Håkansson, and R. G. Nyberg, "Smart parking sensors, technologies and applications for open parking lots: a review," *IET Intelligent Transport Systems*, vol. 12, no. 8, pp. 735–741, 2018.

[5] D. Fernández-Llorca, I. García-Daza, A. Martínez-Hellín, S. Álvarez-Pardo, and M. Á. Sotelo, "Parking assistance system for leaving perpendicular parking lots: Experiments in

daytime\nighttime conditions," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 57–68, 2014.

[6] V. C. Chen, *Inverse Synthetic Aperture Radar Imaging; Principles*. Institution of Engineering and Technology, 2014.

[7] S. Gishkori, D. Wright, L. Daniel, M. Gashinova, and B. Mulgrew, "Imaging moving targets for a forward-scanning automotive sar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 2, pp. 1106–1119, 2019.

[8] J. S. Kulpa, M. Malanowski, D. Gromek, P. Samczynski, K. Kulpa, and A. Gromek, "Experimental results of high-resolution isar imaging of ground-moving vehicles with a stationary fmcw radar," *International Journal of Electronics and Telecommunications*, vol. 59, pp. 293–299, 2013.

[9] C. J. Li and H. Ling, "Wide-angle, ultra-wideband isar imaging of vehicles and drones," *Sensors*, vol. 18, no. 10, p. 3311, 2018.

[10] N. Pandey, G. Duggal, and S. S. Ram, "Database of simulated inverse synthetic aperture radar images for short range automotive radar," in *2020 IEEE International Radar Conference (RADAR)*. IEEE, 2020, pp. 238–243.

[11] N. Pandey and S. S. Ram, "Classification of automotive targets using inverse synthetic aperture radar images," *arXiv preprint arXiv:2101.12535*, 2021.

[12] B. Haywood and R. Evans, "Motion compensation for isar imaging," in *Proceedings of Australian Symposium on Signal processing and applications*, 1989, pp. 112–117.

[13] L. T. Authority, "Handbook on vehicle parking provision in development proposals," 2005.

Autonomous vision-based landing of UAV's on unstructured terrains

1st Evangelos Chatzikalymnios
Electrical and Computer Engineering Department
University of Patras, Greece
chatz@ece.upatras.gr

2nd Konstantinos Moustakas
Electrical and Computer Engineering Department
University of Patras, Greece
moustakas@ece.upatras.gr

***Abstract*—Unmanned Aerial Vehicles (UAVs) technology has enabled the design of many diverse applications in recent years. The development of autonomous landing methods has become a core task, as UAV's navigate in remote and usually unknown environments. In this study we present a vision-based autonomous landing system for UAVs equipped with a stereo camera and an inertial measurement unit (IMU). We utilize stereo processing to acquire the 3D reconstruction of the scene. Next, we evaluate and quantify into map-metrics the factors of the terrain that are crucial for a safe landing. The optimal landing site in terms of flatness, steepness and inclination across the scene is chosen. The pose estimation is obtained by the fusion of stereo ORB-SLAM2 measurements with data from the inertial sensors, assuming no GPS signal. We evaluate the utility of our system using a multifaceted dataset and trials in real-world environments.**

***Index Terms*—unmanned aerial vehicles, vision-based, autonomous landing, stereo processing**

I. Introduction

The use of UAVs has increased in the context of many modern civil applications, including among others delivery, forest surveillance, search and rescue. UAVs equipped with sensors such as optical and thermal cameras, bio-radars and IMUs are life-saving technologies that can provide medical aid to remote environments and support the detection of survivors in cases of emergency.

While navigating in remote environments, UAVs need to be capable of autonomously landing on complex terrains for security, safety and data acquisition reasons. At this point, the study of autonomous landing needs to consider multiple factors and constraints for the development of versatile, robust, and practical autonomous landing systems. Specifically, UAVs often need to operate in GPS-denied or week signal environments. GPS signals are proved to be weak under tunnels, vehicles, forests, metal components, and tall buildings with high density. Furthermore, GPS signal accuracy can be affected by various factors. In particular, weather conditions can quickly reduce signal power while electromagnetic interference such as radio and magnetic fields generate different levels of interference. Therefore, it is essential to study the autonomous landing strategies for drones without GPS signals.

Moreover, when severe failures occur, drones probably lose contact with the ground. In such cases, UAVs should be able to autonomously detect landing sites for an emergency landing. It is crucial to forestall the drone from falling into densely populated areas and protect it from significant damage, using autonomous landing strategies [1], [2].

II. Related work

Autonomous landing has long been standing as an important challenge for UAVs. Vision-based landing in particular, has become attractive because it is passive and does not require any special equipment other than a camera and an on-board processing unit.

Yang et al. [3] propose an autonomous monocular vision-based drone landing system for use in emergencies and unstructured environments. The authors suggest a novel map representation approach that utilizes three-dimensional features extracted from Simultaneous Localization And Mapping (SLAM) to construct a grid-map with different heights. The proposed system gains an understanding of the height distribution of the ground and the obstacle information to a certain extent and subsequently collects the landing area suitable for the UAV. Similarly, Forster et al. [4] use a monocular semi-direct visual odometry (SVO) algorithm to estimate the current UAV's pose given the image stream from the single downward-looking camera. The output of SVO is fused with the data coming from the onboard IMU to estimate the scale of the trajectory. Afterwards, the authors compute depth estimates with a modified version of the Regularized Modular Depth Estimation (REMODE) algorithm [5]. The generated depth maps are then used to incrementally update a 2D robot-centric elevation map [6].

Johnson et al. [7] propose a Lidar-based approach in which an elevation map is computed from Lidar measurements, followed by thresholding the regions based on local slope and roughness of the terrain. Hinzmann et al. [8] present a landing site detection algorithm for autonomous planes. The authors employ a binary random forest classifier to identify grass areas. They extract the most promising landing regions on which hazardous factors such as terrain roughness, slope and the proximity to obstacles are computed to determine the safest landing point. Mittal et al. [9] present a vision-based autonomous landing system for UAV mounted with an IMU and RGB-D camera. The detection algorithm considers several hazardous terrain factors to compute a weighted cost-map based on which dense candidate landing sites are detected. The

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

pose estimation is obtained by fusing IMU, GPS and SLAM data.

A. Contribution

In this work, we employ several terrain factors to determine the safest landing site, using only a stereo camera and an IMU. The 3D reconstruction of the scene is acquired by stereo processing [10] and the pose of the UAV is estimated by fusing raw data from the inertial sensors with the pose obtained from stereo ORB-SLAM2 [11]. We utilize the scene's disparity map and point cloud representation to evaluate the terrain factors and quantify them into map-metrics [12]. These metrics are used to classify the point of the scene and detect candidate landing sites.

More precisely the contribution of our method can be summarized in the following.

- Update of the flatness and steepness map-metrics proposed in [9] and introduction of two novel map-metrics, the inclination and depth-variance. This combination leads to a better perception of the terrain characteristics.
- Landing-decision based on a Bayesian method. We classify the points of the scene into two classes depending on their landing potential. The most promising landing regions are identified and grouped into clusters to later determine the safest landing site.

Finally, we evaluate the performance of our system in different environments such as forest regions with dense vegetation, steep cliffs, stairs and varying altitude, using both a versatile dataset and outdoor trial.

III. Landing Site Detection Algorithm

In this section we present the landing site detection algorithm, including the map-metrics construction, the Bayesian classification of the scene and the final landing sites decision. We ensure that the captured area is perpendicular to the viewing direction by employing the IMU's gyroscope data. When the UAV's yaw, pitch and roll are considerably small the UAV is in hover mode and the autonomous landing procedure commences.

A. Map-metrics Construction

a) Flatness: The flatness of the an area is a property that indicates whether the area is obstacle-free and appropriate for landing. We try to obtain flatness information from the disparity map that we computed considering that the flatness of an area can be represented by the equi-depth region of the map.

By applying a Canny edge detector over the disparity map D, we obtain a binary image $I = Canny(D)$, where non-zero elements represent depth discontinuities. Next, for each pixel $p = (i, j)$ in the image-frame, we compute the distance (in pixels), to nearest non-zero pixel q of the edge-image I [9].

Next the flatness map-metric is calculated as follows:

$$S_{flatness}(p) = min\{euclidean(p, q) \mid I(q) = 1\} \quad (1)$$

b) Depth Variance: The flatness map metric performance is reliant on the Canny filter parameters. Depending on the scene's terrain this map metric may be sensitive to false-positive flatness identification. To smooth out this effect we combine the flatness map metric with a second flatness-oriented metric computed directly from the disparity map, without any pre-possessing step. The depth variance map-metric pixels correspond to the standard deviation of a fixed window with centre the counterpart pixel in the disparity map. High standard deviation values indicates areas with depth discontinuities.

Around each pixel $p = (i, j)$ in the image plane, we apply a fixed window and compute the standard deviation of the included pixel values. The Depth Variance map-metric is calculated as follows:

$$S_{dvar} = e^{-std^2} \quad (2)$$

c) Inclination: Terrain's flatness only, is not adequate to ensure a safe landing procedure. An area with a flat surface can also be so inclined, that landing on it may not be considered as stable or secure. We employ the principal curvatures to determine the inclination of a region under examination. Principal Component Analysis (PCA) is applied on a surface patch of normals to estimate these parameters. We use the same KdTree for the normal and curvature estimation and a curvature score is given to each pixel.

Around each pixel $p = (i, j)$ in the image plane, principal curvature in the z-axis (pc_z) and the corresponding max eigenvalue of curvature (max_c) is computed. Later, the inclination map-metric is calculated as follows:

$$S_{inclination} = \exp(-\frac{\sqrt{|max\_c|}}{\sqrt{|pc\_z|}}) \quad (3)$$

d) Steepness: Another feature of great importance is the steepness of the area around the candidate landing region. We compute the point cloud representation of the scene, we filter from outliers and calculate the point cloud normals. Point cloud normals are a reliable source of information about the steepness of the surface in a specific area.

For each pixel p in the calculated disparity map we find the corresponding point in the generated pointcloud and we compute the angle θ between the normalized surface normal $\hat{n}$ and the z-axis vector in the global frame using the vector dot product. Next the steepness score for each pixel p is the calculated as:

$$S_{steepness} = e^{-\theta^2} \quad (4)$$

B. Bayesian Classification of Landing Points

After evaluating the map-metrics, we perform min-max normalization to scale their values to the same range and remove any biases. At this point the following question arise: Is there any site in the current scene where the UAV can safely land on? To answer this question we classify the scene points into possible and not possible landing sites. We start with the division of the set of scene points viewed in the left stereo image Ω^o into the following two subsets.

1) $\Omega^a \subset \Omega^o$: The set of scene points appropriate for landing
2) $\Omega^n \subset \Omega^o$: The set of scene points not suitable for landing

For this task a Bayesian method is proposed based on the constructed map-metrics and the selection of the following hypothesis.

$$H_a : point\ s\ is\ an\ appropriate\ landing\ site$$
$$H_n : point\ s\ is\ not\ an\ appropriate\ landing\ site$$

Where s in the point of the point cloud corresponding to pixel $n = (i, j)$. According to the Bayes decision test, hypothesis H_a will be selected if $r_a(s, \Omega_a) < r_n(s, \Omega_n)$, where $r_a(s, \Omega_a)$ is the average cost of accepting hypothesis H_a and can be defined as follows :

$$r_a(s, \Omega_a) = \sum_{i=1}^{N} G_{ia} p(s, H_i) = \sum_{i=1}^{N} G_{ia} p(s|H_i) p(H_i) \quad (5)$$

where G_{ia} is the cost of accepting H_a when is true, $p(s, H_i)$ is the mutual probability of s and H_i, and N is the number of the possible landing classes. It is very reasonable to assume that $G_{aa} = G_{nn} = 0$ (zero cost for proper classification) and $K_n * G_{na} = G_{an}$, since erroneous classifications for non-appropriate landing sites is more noxious. Adopting the Maximum-likelihood (ML) criterion and assuming that $p(H_i) = 1/N = 1/2$, (5) is trivially seen to be minimized if the hypothesis H_a is selected when

$$p(s|H_a) > K_n * p(s|H_n) \quad (6)$$

Therefore, all the information regarding the hypothesis selection is inserted in the formula of the probability $p(n|H_i)$. To model this probability we exploit the map-metrics (m) values, calculated in the previous step. Those metrics are considered as the features that we will use to determine the conditional probabilities. Assuming that the features are independent the probabilities can be written as :

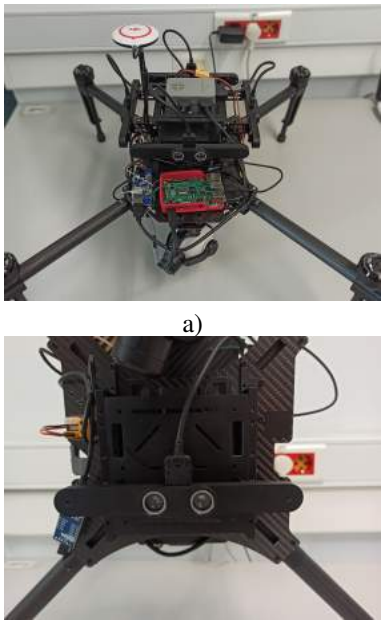
$$p(n|H_i) = \prod_{i=1}^{N_m} f(m_i(n); \lambda) \quad (7)$$

Where, N_m is the number of map-metrics and $f(x; \lambda)$ is the probability density function (pdf) of an exponential distribution. The map-metrics values x are post-normalized to be in the same range [0-1]. Thus, when we evaluate the $p(n|H_n)$ probability we use the x as input. On the other hand, when we evaluate the $p(n|H_a)$, the value 1-x is used.

C. Landing Site Selection

To determine the best landing site, we utilize the set of candidate landing point Ω^a comprised of every potential landing point in image-frame that occurs in (6).

Consequently a k-means clustering algorithm that takes into consideration both candidate's position (X,Y,Z) and normal vector, is applied over the candidate point set to extract the dense landing sites among the scene [13]. The centroid of the biggest cluster is considered as the safest landing site and the corresponding point in the pointcloud is identified. Finally, we compute the dominant clusters area and compare it with the UAV size.



a)

b)

Fig. 1. Drone setup. a) Raspberry Pi connection with Flight controller, Guidance core and DC-DC converter b) Used down-looking stereo camera.

IV. Experimental Evaluation

We evaluate the utility of our system using a versatile dataset and extensive experiments in real-world environments. Our dataset is comprised of 100 grayscale stereo-image pairs taken from the down-looking stereo camera. The captured terrains include scenes of road sections,vehicles, trees, dense vegetation and buildings (Fig. 2). Moreover, the dataset was organised in different classes depending on the UAV's altitude (0-30m).

Fig. 2. A sample of the used dataset.

We use DJI Matrice 100 with mounted DJI guidance system. Guidance's IMU accelerometer and gyroscope data are

fused with ORB_SLAM's measurements for the UAV pose estimation. The stereo camera underneath the UAV produces grayscale images with a resolution of 320 × 240 pixels at 20Hz. Robotic Operating System (ROS) is used as the middleware on the Raspberry Pi 4 which runs all the processes for state estimation, sensor fusion and landing site detection. Furthermore, we exploit a feature-based method ORB_SLAM2 to estimate the UAV's pose and attain the fixed point in three-dimensional space. Additionally, we utilize IMUs gyroscope and accelerometer information, which is fused with position and orientation measurements from SLAM using an unscented Kalman filter [14], [15].

A. Map-Metrics Evaluation

In Fig. 3 we demonstrate the performance of the landing site detection algorithm as a function of the UAV's altitude.

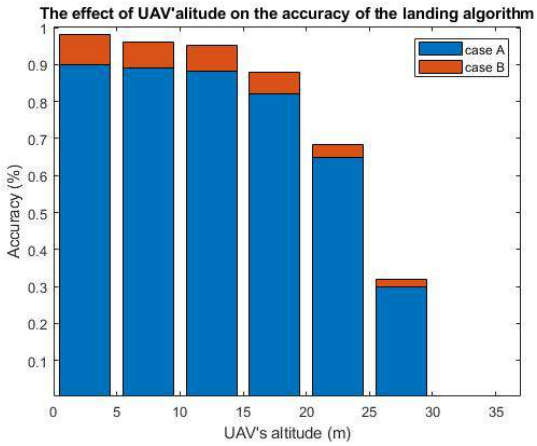


Fig. 3. The accuracy of safe landing site detection as a function of the UAV's altitude. In case A (blue), the flatness and the steepness map-metrics are utilized. In case B (orange) we show the accuracy improvement by adding depth-variance and inclination map metrics.

The performance is evaluates by means of accuracy. For every image in the dataset the algorithm detects a landing site. The accuracy is defined as the proportion of the safe sites detected by the algorithm across all images in the dataset. As expected, the performance decreases with altitude however the algorithm performs well for altitudes less or equal than 15 meters (accuracy over 96%).

Table 1 illustrates the time and memory consumption of every single metric and process in the system. The results indicate that the flatness map-metric is the most time-demanding procedure. The time consumed is highly affected by the distance transform operation, which varies with the image content of the binary map obtained from the canny edge operation. However, the flatness map-metric is highly informative and performs very efficiently in terrains with uniform and wide flat areas. One crucial parameter is the canny edge detector sensitivity. Depending on the scene's terrain this map metric may be sensitive to false-positive flatness identification. We counterbalance this effect by utilizing the the depth-variance map metric. The depth-variance map metric performs pixel-wise calculations and thus it runs a lot faster. It can be considered as less informative than the flatness map-metric however, it is less sensitive to false-positive identifications. It is a fast calculation that gives a big penalty to pixels that are part of regions with high non-flat probability.

TABLE I
RUN-TIME AND MEMORY CONSUMPTION FOR PROCESSING EACH FRAME. EVALUATED ON RASPBERRY PI 4, BROADCOM BCM2711, QUAD CORE CORTEX-A72 (ARM V8) 64-BIT SOC @ 1.5GHZ

Table Column Head		
Algorithm	***Time Cost (ms)***	***Memory (MB)***
Disparity map estimation	12 ± 2.6	19.7 ± 3.5
Point Cloud creation	5 ± 1.2	17.5 ± 3.2
Flatness map-metric	240 ± 2.9	22.2 ± 6.8
Steepness map-metric	95 ± 2.5	69.3 ± 0.1
Inclination map-metric	55 ± 3.2	20.9 ± 2.3
Depth Variance map-metric	16.5 ± 3.8	18.7 ± 2.5
Bayesian Classification	28.8 ± 1.6	13.1 ± 0.2
Clustering	2 ± 0.7	20.7 ± 6.7
Total	***454.3 ± 18.5(ms)***	***202.1 ± 25.3 (MB)***

The Steepness map metric utilizes the point cloud normal vectors to locally compute the steepness of the scene. It is a time efficient map-metric and performs great in identifying true-positive, non-steep sites. The inclination map metric is used the gain information about the inclination of a wider area. The results indicate great performance in identifying true-positive, incline sites. At the same time, it gives a high score in areas with a very high probability of being non-incline which makes it highly valuable and improves the performance of the landing site detection.

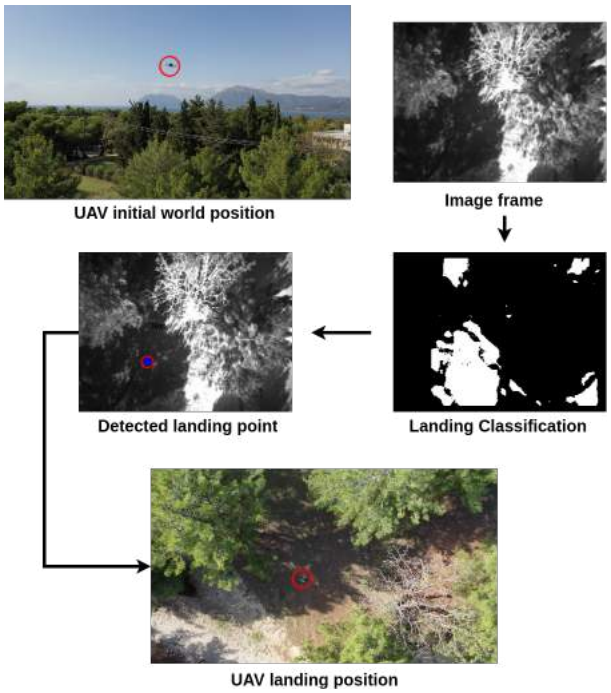


Fig. 4. Illustrations of real world experiment in dense wooded area.

B. Outdoor Trials

In this section we evaluate the utility of our system extensive experiments in real world. The proposed landing method is described next. When emergency landing is needed the following method is designed. The UAV is set to hover mode, and the scene beneath the UAV is captured. The Navigation node continuously checks if the UAV hovers and updates the Boolean message of the hover ROS topic. The Landing algorithm node subscribes to the ROS images topics, and the hover topic and begins the process when hover mode is set on. The landing algorithm outputs the best landing site. The landing sites world coordinates are calculated taking into account the UAV's pose in world coordinates (X,Y,Z) given from IMU-SLAM fusion and the detected landing point's Image coordinates (x,y,z). Afterwards, landing is taking place, and the procedure comes to an end.

Our landing site algorithm safely detects landing sites in various scenarios. The tested environments come form the same distribution as the dataset evaluated in section IV-A, while the outdoor trials follow the predictions of our algorithm (Fig. 3). In Fig. 4 we demonstrate the outcome of an outdoor trial in a dense wooded area.

V. Conclusion

In this work, a vision-based autonomous landing system for UAVs equipped with a down-looking stereo camera and an IMU is presented. Our landing site detection algorithm considers several terrain factors including the flatness, inclination, steepness and depth-variance that are quantified into map-metrics. Subsequently, we employ a Bayes-based method to classify the points of the scene based on their landing appropriateness and detect candidate landing sites.

The system's performance is proven to be affected by the operation altitude however, the accuracy of the algorithm is over 96% for an altitude less or equal to 15 meters. The proposed system is computationally efficient as it runs online on a Raspberry Pi 4 embedded computer with other processes for state-estimation and bio radar processing being run in the background. The utility of our system is demonstrated using extensive real-world environment experiments.

Future research should consider the potential effect of water surface more carefully. For example, lakes and sea are uniform, low-contrast terrains that the algorithm may false identify as potential landing sites.

References

[1] Liu Yubo et al. "Survey of UAV Autonomous Landing Based on Vision Processing, Advances in Intelligent Networking and Collaborative Systems," Springer International Publishing, pages pages 300-311, 2021.

[2] Kanellakis, Christoforos, and George Nikolakopoulos. "Survey on computer vision for UAVs: Current developments and trends." Journal of Intelligent & Robotic Systems 87.1 (2017): 141-168.

[3] T. Yang, P. Li, H. Zhang, J. Li, and Z. Li, "Monocular vision slam-based uav autonomous landing in emergencies and unknown environments,"Electronics, vol. 7, no. 5, p. 73, 2018.

[4] C. Forster, M. Faessler, F. Fontana, M. Werlberger, and D. Scaramuzza, "Continuous on-board monocular-vision-based elevation mapping applied to autonomous landing of micro aerial vehicles," in2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE,2015, pp. 111–118.

[5] M.Pizzoli, C.Forster, and D.Scaramuzza,"Remode: Probabilistic, monocular dense reconstruction in real time," in2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014, pp. 2609–2616.

[6] P. Fankhauser, M. Bloesch, C. Gehring, M. Hutter, and R. Siegwart, "Robot-centric elevation mapping with uncertainty estimates," in Mobile Service Robotics. World Scientific, 2014, pp.433–440.

[7] A. E Johnson et al. "Lidar-based hazard avoidance for safe landing on mars." JGCD, 25(6):1091–1099, 2002.

[8] T. Hinzmann, T. Stastny, C. Cadena, R. Siegwart and I. Gilitschenski, "Free LSD: Prior-Free Visual Landing Site Detection for Autonomous Planes," in IEEE Robotics and Automation Letters, vol. 3, no. 3, pp. 2545-2552, July 2018, doi: 10.1109/LRA.2018.2809962.

[9] M. Mittal, A. Valada, and W. Burgard, "Vision-based autonomous landing in catastrophe-struck environments," arXiv preprint arXiv:1809.05700, 2018.

[10] Wöhler, Christian. " 3D computer vision: efficient methods and applications." Springer Science & Business Media, 2012.

[11] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," in IEEE Transactions on Robotics, vol. 33, no. 5, pp. 1255-1262, Oct. 2017, doi: 10.1109/TRO.2017.2705103.

[12] F. Bonin-Font,A. Ortiz & G. Oliver, "Visual Navigation for Mobile Robots: A Survey," Journal of Intelligent & Robotic Systems 53, 263 (2008).

[13] Point Cloud Library, Conditional Euclidean Clustering [source code]. https://pointclouds.org/documentation/tutorials/cluster_extraction.html.

[14] Simon J. Julier, Jeffrey K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI, (28 July 1997); https://doi.org/10.1117/12.280797

[15] EE. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373), Lake Louise, AB, Canada, 2000, pp. 153-158, doi: 10.1109/ASSPCC.2000.882463.

GESTURE LEARNING FOR SELF-DRIVING CARS

Ethan Shaotran, Jonathan J. Cruz, and Vijay Janapa Reddi

Harvard University
eshaotran@college.harvard.edu

ABSTRACT

Human-computer interaction (HCI) is crucial for safety as autonomous vehicles (AVs) become commonplace. Yet, little effort has been put toward ensuring that AVs understand human communications on the road. In this paper, we present Gesture Learning for Advanced Driver Assistance Systems (GLADAS), a deep learning-based self-driving car hand gesture recognition system developed and evaluated using virtual simulation. We focus on gestures as they are a natural and common way for pedestrians to interact with drivers. We challenge the system to perform in typical, everyday driving interactions with humans. Our results provide a baseline performance of 94.56% accuracy and 85.91% F1 score, promising statistics that surpass human performance and motivate the need for further research into human-AV interaction.

***Index Terms*—** Human-Computer Interaction, Autonomous Vehicles, Deep Learning, Hand Gestures

1. INTRODUCTION

Autonomous vehicles (AVs) are anticipated to improve transportation efficiency, ease, and costs. Realizing these benefits will require AVs, and more specifically Self-Driving Cars (SDCs), to feature trustworthy human-computer interaction skills [1, 2]. However, little progress has been made towards human-SDC interaction in human-populated driving environments. Situations in which an SDC and pedestrian must negotiate right of way at an intersection, for example, will require the car to interact and operate accordingly.

While a multitude of studies investigate car-to-pedestrian communication [3, 4, 5], the conveyance of an SDC's intent to a pedestrian, we instead research a communication mode of the less-explored pedestrian-to-car communication, the reception of a pedestrian's intent by an SDC. In particular, we focus on hand gestures, which Vinkhuyzen and Cefkin [3] find to be a customary and universal practice in negotiating right of way—for example, a pedestrian may give up their right of way by gesturing to the car to proceed first at an intersection. Organizations such as the United States Government [6] have stated the need for future SDCs to recognize hand gestures—yet, they are lightly studied, especially in the realm of human-SDC interaction.

In this paper, we describe a new end-to-end methodology for enabling SDCs to interact with pedestrians using hand gestures. We develop GLADAS, an SDC hand gesture recognition system developed and evaluated using a virtual simulation with common real-life driving scenarios and human-SDC interactions. The system is composed of an efficient two-model recognition algorithm that identifies pedestrians and classifies any potential gestures they make. We then run 28,000 simulation tests of various interaction scenarios to evaluate and benchmark GLADAS's performance.

Results from our evaluation methodology provide an accuracy of 94.56% and F1 score of 85.91%. In the context of an SDC, these results serve as an initial baseline for future recognition systems built and evaluated using this methodology. For reference, to our knowledge, the current only metric on human classification of hand gestures is 88.4% accuracy by NVIDIA [7]. Our results, the first of their kind, set a baseline for future work to improve on and quantitatively motivate the need for more research on human-AV interaction.

In the remainder of the paper, we debrief related work (Section 2) and describe the components of the simulation (Section 3). Then, we explain the GLADAS recognition system (Section 4) and measure its performance (Section 5). Finally, we discuss the findings and their implications on the future of Human-Computer Interaction for SDCs (Section 6) and conclude the paper (Section 7).

2. RELATED WORKS

Research on pedestrian-to-car communication can be split into two distinct categories. The majority focus of related work is on pedestrians' *implicit communication*, in which intent is inferred from a human's indirect actions—for example, head orientation [8] and kinematic walking trajectory [9]. In contrast, the basis of our work is *explicit communication*, in which the pedestrian directly conveys their intent to the SDC.

The social relevancy of hand gestures in roadway scenarios is widey noted. Rasouli and Tsotsos [10] find that hand gestures are prominent forms of explicit communication that pedestrians use to communicate with cars. However, only a limited number of studies have been conducted on human hand gesture recognition in the roadway setting. Tao and Ben [11] developed an accelerometer, placed at an intersection, to

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

classify Chinese policepeople's gestures and mirror the command in the above traffic lights. Guo et al. [12] employed statistical techniques and the nearest neighbor classifier for recognizing gestures in still, staged images of Chinese policepeople taken by normal cameras. Their results indicated accuracies between 60% and 100% for each hand gesture type.

It is evident that human-SDC interaction is a poorly explored problem, especially in the context of a *self-driving car* understanding pedestrians and other road users' hand gestures. A comprehensive effort to classify gestures of all people, not only, for example, policepeople, in roadway scenarios is needed. Additionally, gesture recognition should occur from the first-person view of an SDC—in line with how today's SDCs perceive their environment. GLADAS, which classifies and reacts to hand gestures in SDC sensors' real-time video streams, seeks to fill these gaps.

3. VIRTUAL SIMULATION

Our research features a controlled simulation with human pedestrians. Our rationale for using a simulator is simple: it provides public safety, realism, and ease for experimentation.

- **Safety.** Our simulation enables rapid and repeated testing with minimized risk to the general public [13] and no immediate real-world consequences. Real-life testing would require closing off intersections and placing real humans at each, potentially hazardous if an SDC acts on a misclassified gesture.
- **Realism.** Research shows that simulated SDC data can be successfully used for training real-world SDCs [14]. SDC simulators such as Intel's CARLA [15] use comparable 3D human and simulation models for testing.
- **Ease of Experimentation.** We are able to adjust factors such as the surrounding objects, road layout, and pedestrian positioning. This enables a wide range of driving scenarios, in which an exact scenario can be constructed for testing.

To achieve the aforementioned objectives, the simulation system is composed from three main parts: the Simulated Environment, AirSim Self-Driving Car, and Image Streamer.

3.1. Simulated Environment

We build our simulation in Unreal Engine 4. To make our simulation as representative of real-world driving as possible, we add houses, parked cars, roads, road signs, and vegetation, shown in Figure 1(a). We also add clothed pedestrians with animated gestures. The commands associated with each gesture, as suggested by Gupta et al. [16], are the four most commonly used commands in road driving: "Go Forward," "Stop," "Go Right," and "Go Left" (along with a base class "No Gesture"). A human gesturing to a car, viewed from within the car, is shown in Figure 1(b).

(a) Street view of the simulator.

(b) View of a pedestrian from inside the SDC.

Fig. 1. Environment used to simulate real-world scenarios.

3.2. AirSim Self-Driving Car

We use the AirSim [17] Unreal plugin to simulate the SDC itself. We stream real-time video data from the SDC's RGB camera using the Image Streamer mechanism described below, and the driving decisions returned from the gesture recognition system direct the car to move accordingly.

3.3. Image Streamer

Frames from the SDC camera are streamed in real-time to a Python client. We observed an inverse relationship between the spatial quality (i.e. image resolution and size) and temporal quality (i.e. frame rate or Frames Per Second (FPS)) of the frames streamed. Spatial quality allows an algorithm to view the pedestrian in detail in *each* frame, while temporal quality views the gesture motion in detail throughout *all* frames.

Real AVs process data between 10 and 40 FPS [18], depending on vehicle speed. In our case the car is halted when it analyzes a gesture and decides to move, such as at a stop sign or crosswalk. Thus, a frame rate above 10 FPS is sufficient.

We optimize and balance the aforementioned trade-off by streaming frames of size `1280x480 px` at a rate of 12.62 FPS. This involves setting the simulator clock speed to 0.14 (1 real second equals 0.14 simulator seconds), allowing the computer more time to get frames per simulator second. This ratio can be further adjusted to increase FPS, but at the cost of training time. Under this configuration, frames are sent to the gesture recognition system, detailed in the next section.

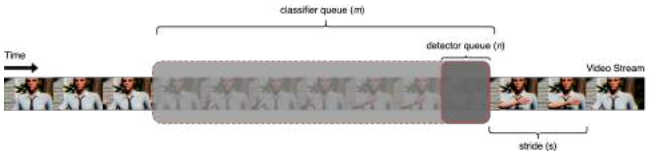


Fig. 2. The structure of the proposed two-model architecture. As frames are streamed in real-time, the lightweight detector processes every s-th frame, where s is the stride and detector queue (n) is 1 frame. The detector acts as a trigger for the classifier, which operates on the next incoming m frames.

4. RECOGNITION SYSTEM METHODOLOGY

Our real-time gesture recognition system's underlying algorithm integrates two models, cascaded sequentially using a sliding window approach. The first, a lightweight Pedestrian Detector (PD) model, detects for pedestrians in every fifth ($s = 5$) frame as they are streamed in real-time. The PD acts as a trigger; if it detects a pedestrian in the most recent frame (i.e. in the SDC's real-time field of view), it activates the more time and computing power-intensive Gesture Classifier (GC) model, which classifies the pedestrian's gesture from the next m = 40 incoming frames of the video stream. This process is illustrated in Figure 2, and each model is detailed below.

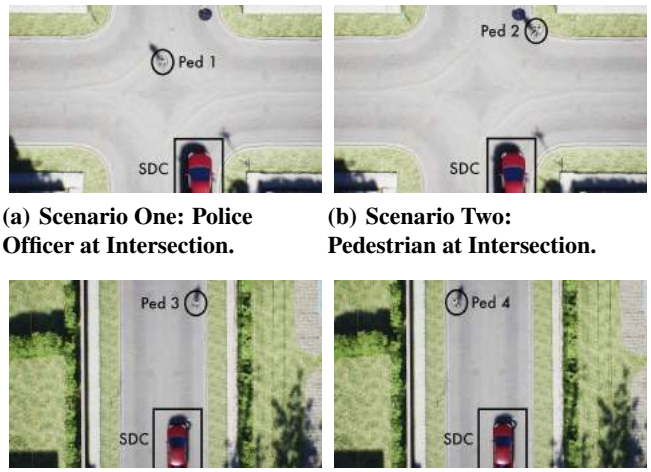
4.1. Pedestrian Detector (PD)

The PD detects any pedestrians in the SDC's field of view and returns the coordinates of their upper body. It is designed to be lightweight and fast, so it can be applied in real-time without incurring heavy computational costs.

The PD uses OpenCV's Pedestrian Recognition model with the Histogram of Oriented Gradients (HOGS) [19] method. To increase speed, we scale the `1280x480 px` frame to `768x288 px` for processing, and re-map the outputted bounding box coordinates of any detected pedestrians back to the original frame size. Finally, we perform additional cropping of the coordinates—1/7 from the top, 1/3 from the bottom, 1/9 from the left, and 1/5 from the right—which we found best isolated a pedestrian's upper body region for optimal gesture classification.

4.2. Gesture Classifier (GC)

Upon the PD's detection of a pedestrian in the frame, the GC captures and crops the next 40 frames using the upper body bounding box coordinates. We choose 40 frames, as it takes ∼40 frames to capture the entire motion of the pedestrian's gesture within the simulator. We then perform a temporal transform (select a random sample of 32 consecutive frames) and a spatial transform (resize the image frames to `112x112 px`) to meet the model's input requirements.



(a) Scenario One: Police Officer at Intersection.

(b) Scenario Two: Pedestrian at Intersection.

(c) Scenario Three: Pedestrian Crossing from Right.

(d) Scenario Four: Pedestrian Crossing from Left.

Fig. 3. The Self-Driving Car and human are positioned as shown for each interaction scenario.

We input these images into a deep learning gesture recognition model by Köpüklü et al. [20]. It is a 3D Convolutional Neural Network (3D CNN) pretrained on the 20BN-Jester Dataset's [21] *training* split—118,562 real-life videos of humans enacting hand gestures. There are 27 different gesture classes—our five target hand gestures ("Go Forward," "Stop," "Go Right," "Go Left," and "No Gesture"), along with 22 others with no immediate relevancy to pedestrian-to-car communication (e.g. "Drumming Fingers").

A 3D CNN processes chronological sequences of images, necessary to preserve the temporal semantics of the action. For example, the full motion of a hand waving may signify "Hello," while a single frame might instead appear as "Stop."

The model returns 27 class confidences. We remove the extraneous classes and normalize the five main classes' confidence scores; the gesture with the highest score is the predicted gesture. The SDC can then react and drive as directed.

5. EXPERIMENTAL SETUP

We set up several SDC-Pedestrian interaction scenarios, which were carefully chosen as models of typical, real-life driving scenarios [22].

1. Police Officer-Controlled 4-Way Intersection (Figure 3(a))
2. Pedestrian Crossing at 4-Way Intersection (Figure 3(b))
3. Pedestrian Crossing from Right at Mid-Block (Figure 3(c))
4. Pedestrian Crossing from Left at Mid-Block (Figure 3(d))

In Scenario 1, the officer performs all five gestures. In Scenarios 2-4, the pedestrian uses "Go Forward," "Stop," and "No Gesture"; the other gestures are generally inapplicable. Each scenario varies in lighting, surrounding objects, scenery, positioning, and SDC-pedestrian distance. Each gesture in each scenario is performed 2,000 times, totalling 28,000 tests.

6. RESULTS

6.1. Accuracy and F1 Analysis

To evaluate performance, we calculate the accuracy and F1 score of the predictions for each gesture in each scenario. While accuracy is a more intuitive metric, F1 score additionally penalizes False Negatives (FNs) and False Positives (FPs)—in the context of SDCs, these sorts of misclassifications could lead to the car acting incorrectly.

The accuracy and F1 score (with a classification confidence threshold of 0.40) over all scenarios and gestures are shown in Table 1 and Table 2, respectively. We make a couple initial observations. First, recognition of gestures across all scenarios is better than a baseline random guess. Second, recognition of "Go Right" in Scenario One is relatively poor—the data collected suggests the GC regularly misidentified "Go Right" gestures as "Go Left," perhaps due to their similar hand motions. Lastly, recognition of "No Gesture" is better than most other gestures for each scenario.

Altogether, the results yield an average accuracy of 94.56% and average F1 score of 85.91%. That the accuracy is larger than the F1 score suggests the notable presence of FPs and FNs—considering only accuracy might provide an optimistic perception of performance. In other words, judging from the F1 score, the true performance may be a bit lower than the accuracy of 94.56%.

For comparison, NVIDIA [7] found that human state-of-the-art classification accuracy on the NVIDIA Gesture Dataset, a well-known alternative to Jester, was 88.4% (to our knowledge, currently the only metric of human classification of hand gestures). Additionally, our accuracy scores are similar to Guo et al. [12] but roughly 20% better for "Go Right" and "Go Left." Overall, in the context of roadway driving, maximizing accuracy and F1 score is crucial, as misclassifications could result in serious roadway accidents.

6.2. Simulation vs. Reality

To demonstrate the gap between GLADAS's performance in simulation versus reality (sim-to-real), we also run its GC model on real-life data: the Jester dataset's *validation* split of 2,519 real videos of humans making the five relevant gestures. The average accuracy and F1 score across the gestures are 99.67% and 99.14%, respectively. The difference in classification performance on real vs. simulated data (99.67% vs. 94.56% for accuracy and 99.14% vs. 85.91% for F1 score) helps quantify the sim-to-real gap—a likely consequence of the difficulty of perfectly simulating human hand gestures and demeanor. Notably, however, our simulation evaluation methodology does not provide overly confident estimates of real-world performance, but rather more conservative ones.

		Gesture				
		Go Straight	***Stop***	***No Gesture***	***Go Right***	***Go Left***
Scenario	***1***	93.4%	95.6%	93.3%	88.2%	95.0%
	2	94.9%	93.3%	98.1%	n/a	n/a
	3	93.6%	91.2%	96.4%	n/a	n/a
	4	92.3%	98.7%	99.9%	n/a	n/a

Table 1. The accuracy of each Scenario-Gesture pairing.

		Gesture				
		Go Straight	***Stop***	***No Gesture***	***Go Right***	***Go Left***
Scenario	***1***	76.7%	84.9%	83.2%	62.6%	83.1%
	2	88.8%	87.0%	96.4%	n/a	n/a
	3	85.4%	82.2%	93.3%	n/a	n/a
	4	81.9%	97.5%	99.7%	n/a	n/a

Table 2. The F1 score of each Scenario-Gesture pairing.

6.3. Implications

In order to deploy SDC gesture recognition systems that save lives, detection, classification, and reaction to human hand gestures must be trustworthy and reliable. As such, the main goal of our work was to further bridge the connection gap between SDCs and humans. Our results show that SDCs can understand gestures with a 94.56% accuracy and 85.91% F1 score using this system. Although exceeding current metrics of human accuracy, SDCs will be held to a higher standard—thus, these results strongly enforce the need for continued research and development of more reliable algorithms.

Our methodology presents one such way to do so: develop and scrutinize a recognition system in a simulated environment with common SDC-pedestrian interactions. Future work could incorporate components such as image segmentation and specialized hardware to improve upon our progress.

7. CONCLUSION

We presented GLADAS, a hand gesture recognition system designed to bridge the gap in human-SDC interaction. Using a virtual simulation based on real life driving, we design and evaluate the recognition system, which is composed of a pedestrian detector and gesture classifier. We challenge the system to recognize a human pedestrian's hand gesture in four common human-populated roadway scenarios. The results provided by our evaluation methodology motivate continued SDC-pedestrian interaction research, particularly within the context of gesture recognition and trustworthy interaction systems. We hope GLADAS inspires further research into SDC gesture recognition, paving the way to full autonomy.

8. REFERENCES

[1] Morteza Lahijanian and Marta Kwiatkowska, "Social trust: A major challenge for the future of autonomous systems," in *AAAI Fall Symposium Series*, Sep 2016.

[2] Amir Rasouli and John K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *arXiv:1805.11773 [cs]*, May 2018.

[3] Erik Vinkhuyzen and Melissa Cefkin, "Developing socially acceptable autonomous vehicles," *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2016, no. 1, pp. 522–534, 2016.

[4] Karthik Mahadevan, Sowmya Somanath, and Ehud Sharlin, "Communicating awareness and intent in autonomous vehicle-pedestrian interaction," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, CHI '18, pp. 429:1–429:12, ACM, event-place: Montreal QC, Canada.

[5] Milecia Matthews, Girish Chowdhary, and Emily Kieson, "Intent communication between autonomous vehicles and pedestrians," *arXiv:1708.07123 [cs]*, Aug 2017.

[6] U.S. Senate, *Hands Off: The Future of Self-Driving Cars*, U.S. Government Publishing Office, Mar 2016.

[7] Pavlo Molchanov, Xiaodong Yang, Shalini Gupta, Kihwan Kim, Stephen Tyree, and Jan Kautz, "Online detection and classification of dynamic hand gestures with recurrent 3d convolutional neural networks," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016, p. 4207–4215.

[8] Amir Rasouli, Iuliia Kotseruba, and John K. Tsotsos, "Understanding pedestrian behavior in complex traffic scenes," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 61–70, Mar 2018.

[9] Sujeong Kim, Stephen Guy, Wenxi Liu, Rynson Lau, Ming Lin, and Dinesh Manocha, "Predicting pedestrian trajectories using velocity-space reasoning," *International Journal of Robotics Research*, vol. 34, Jan 2014.

[10] Amir Rasouli, Iuliia Kotseruba, and John K. Tsotsos, "Agreeing to cross: How drivers and pedestrians communicate," *arXiv:1702.03555 [cs]*, Feb 2017.

[11] Yuan Tao and Wang Ben, "Accelerometer-based chinese traffic police gesture recognition system," *Chinese Journal of Electronics*, vol. 19, no. 2, Apr 2021.

[12] Fan Guo, Jin Tang, and Changjun Zhu, "Gesture recognition for chinese traffic police," in *2015 International Conference on Virtual Reality and Visualization (ICVRV)*, Oct 2015, pp. 64–67.

[13] Matthew O'Kelly, Aman Sinha, Hongseok Namkoong, John Duchi, and Russ Tedrake, "Scalable end-to-end autonomous vehicle testing via rare-event simulation," in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, Dec 2018, p. 9849–9860.

[14] Matthew Johnson-Roberson, Charles Barto, Rounak Mehta, Sharath Nittur Sridhar, Karl Rosaen, and Ram Vasudevan, "Driving in the matrix: Can virtual worlds replace human-generated annotations for real world tasks?," *arXiv:1610.01983 [cs]*, Oct 2016.

[15] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun, "Carla: An open urban driving simulator," *arXiv:1711.03938 [cs]*, Nov 2017.

[16] Surabhi Gupta, Maria Vasardani, and Stephan Winter, "Conventionalized gestures for the interaction of people in traffic with autonomous vehicles," in *Proceedings of the 9th ACM SIGSPATIAL International Workshop on Computational Transportation Science*, Oct 2016, IWCTS '16, p. 55–60.

[17] Shital Shah, Debadeepta Dey, Chris Lovett, and Ashish Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Robotics*, 2018, p. 621–635.

[18] Ming Yang, Shige Wang, Joshua Bakita, Thanh Vu, F. Donelson Smith, James H. Anderson, and Jan-Michael Frahm, "Re-thinking cnn frameworks for time-sensitive autonomous-driving applications: Addressing an industrial challenge," in *2019 IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, Apr 2019, p. 305–317.

[19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2005, vol. 1, pp. 886–893.

[20] Okan Köpüklü, Ahmet Gunduz, Neslihan Kose, and Gerhard Rigoll, "Real-time hand gesture detection and classification using convolutional neural networks," *arXiv:1901.10323 [cs]*, Jan 2019.

[21] Joanna Materzynska, Guillaume Berger, Ingo Bax, and Roland Memisevic, "The jester dataset: A large-scale video dataset of human gestures," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. Oct 2019, p. 2874–2882, IEEE.

[22] Susan T. Chrysler, Omar Ahmad, and Chris W. Schwarz, "Creating pedestrian crash scenarios in a driving simulator environment," *Traffic Injury Prevention*, vol. 16 Suppl 1, pp. S12–17, 2015.

Collision Prediction using UWB and Inertial Sensing: Experimental Evaluation

Aarti Singh
Electrical and Systems Engineering
Washington University in St. Louis, USA
aartisingh@wustl.edu

Neal Patwari
McKelvey School of Engineering
Washington University in St. Louis, USA
npatwari@wustl.edu

***Abstract*—Real-time proximity and collision detection via radio frequency (RF) distance measurements has application in smart helmets, drones, autonomous vehicles, and social distancing. In this paper we evaluate ACED, a range-based, infrastructure-free, distributed algorithm that utilizes inter-node range data and intra-node acceleration data to estimate the recent relative positions of each node and to predict impending collisions between any pair of nodes. The framework is tested and validated using experimental data from a testbed of mobile nodes which use ultra-wideband (UWB) ranging and inertial sensing. ACED is shown to outperform two state-of-the-art methods.**

***Index Terms*—Collision prediction, multidimensional scaling, autonomous vehicles**

I. Introduction

Autonomous and real-time collision prediction and collision avoidance is crucial in a world filled with multiple mobile entities operating in the close vicinity of each other. Collisions, which do happen [1], are both life-threatening and expensive. GPS and lidar are insufficient to reliably predict the collision of small objects moving quickly towards each other, e.g., multiple drones, or a smart helmet and a baseball. In many cases, it will be possible to add a radio frequency (RF) tag to nodes, vehicles, or objects that need to monitor to avoid collisions. However, RF tag localization systems are insufficient to predict collisions because they do not predict future positions, and they require a fixed, known-location infrastructure to calculate global map of nodes' locations, which may not be present, inconvenient, or expensive to deploy for the application. We argue that, fundamentally, collision prediction from range measurements should be distributed, local, and relative. Fortunately, collision between two objects does not require the coordinates of each object in a global coordinate system because the collision between two objects is a matter of their relative kinematics, such as their relative position, velocity, and acceleration. In this paper, we present a method to address this gap by enabling mobile agents of any size or speed to predict impending collisions without relying on a centralized infrastructure or a global reference.

A popular approach to obtain relative positions is multidimensional scaling (MDS). In MDS, 'dissimilarities' between each pair of objects are mapped into a low dimensional relative coordinates such that the distances between nodes are preserved as much as possible. MDS has been utilized in the localization research extensively [2], however, MDS is a centralized algorithm as it requires all dissimilarities to be known by one processing unit. For N nodes, classical MDS has a computational complexity of $O(N^3)$. A distributed method of estimating location is proposed in [3], is implemented with known-location infrastructure nodes. Based on the same work, another approach is presented to obtain a relative map of objects in motion, which although does not require known-location infrastructure, but is centralized in its implementation [4]. Another challenge with using MDS to generate relative kinematics over a time period is that, since there is no fixed frame of reference, the generated map can undergo random translation, rotation, and flip. Therefore, without infrastructure, successive applications of MDS over time provides incorrect kinematics. A modification of classical MDS such that a common frame of reference is maintained for position and higher order kinematics (velocity and acceleration) is implemented in [5]. The relative kinematics are estimated using higher order derivatives of squared distance measurements. However, this method is highly sensitive to noise in range measurements. In order to predict collision from RF range measurements we need relative kinematics estimators which are tolerant to noisy measurements. One extension of this work is implemented to produce a kinematics based collision prediction model, but it is limited to only linear motion [6]. We present a robust, decentralized, infrastructure-less algorithm 'Autonomous Collision Estimation for Dynamic Motion' (ACED), that produces quality relative kinematics of moving objects by using noisy inter-node range measurements and intra-node acceleration data. With the estimated kinematics, this distributed scheme predicts the future trajectory and any impending collision without requiring known-location devices. In addition, our solution can be complementary to other widely adopted technologies for collision prediction. These technologies involving either active sensors such as lidar [7] or radar [8], or passive sensors [9] such as cameras, are dependent on size, shape, reflective properties, luminescence of, and distance between the objects involved in the collision. The algorithm presented in this paper is free of such limitations involving physical properties of the objects.

II. Problem Statement

We take a network of N unknown-location nodes in a D-dimensional space. Over a period of time, node i collects

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

pairwise range measurements between itself and its neighbors $\mathcal{N}_i$. We denote the range measurement between node i and $j \in \mathcal{N}_i$ at time t as $\delta_{i,j}^t$. A real-world scenario is considered in which nodes move with arbitrary motion. The problems we explore include:

1) estimation of the coordinates $\boldsymbol{x}_i^t$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where $\boldsymbol{x}_i \in \mathcal{R}^D$ given pairwise range measurements $\{\delta_{i,j}^t\}$ and individual acceleration measurements, $\boldsymbol{\alpha}_i^t$, both taken over the time window $t = 1, \ldots, T$;
2) predicting an impending collision between i and any neighbor node in $\mathcal{N}_i$ at a time soon after $t = T$.

We assume that the primary goal of the system is to predict collisions, but in the case of an impending collision, the recent positions may be useful for the system reaction, for example, to know which direction to swerve.

III. Proposed Algorithm

To achieve the goals we articulate in the Introduction, in this section, we formulate a new cost function based on errors between measured and calculated ranges and acceleration over time for all nodes. Our insight is that distributed tracking can be formulated in a distributed, low computational complexity manner by using a majorization framework. In this framework, each device estimates its recent positions locally by minimizing its local cost function and broadcasting its newest position estimates to its neighbors. Each node successively refines its recent position estimates based on its range and acceleration measurements in addition to the most recent received position estimates from its neighbors. The majorization approach ensures non-increasing local cost functions. Since these local costs contribute additively to the global cost, thus, the global cost function will be non-increasing. Further, the local optimization step is low complexity because it is based on finding the minimum of a quadratic (majorizing) expression. Finally, the distributed optimization is guaranteed to converge, because the majorization approach guarantees each round's global cost is non-increasing. Our algorithm follows similarly to the *scaling by majorizing a complicated function* (SMACOF) [10] approach, but expands SMACOF to enable simultaneous estimation of multiple recent positions, and to enable use of acceleration measurements in the cost function.

A. Proposed Cost Function

As described in Section II, our problem is to estimate node positions $X = \{\boldsymbol{x}_i^t\}_{i,t}$ to match measured ranges $\{\delta_{i,j}^t\}$ and node acceleration measurements $\{\boldsymbol{\alpha}_i\}$. Our cost function S penalizes any coordinates that increase the squared error. We divide S into components for each node, $S(X) = \sum_i S_i(X)$, where the local cost function $S_i(X)$ is:

$$\sum_{t=1}^{T} \left[\sum_{j \in \mathcal{N}_i} w_{i,j}^t \left[\delta_{i,j}^t - d_{i,j}(X) \right]^2 + r_i \left[\boldsymbol{\alpha}_i^t - \boldsymbol{a}_i(X) \right]^2 \right],$$

where the first term represents the error between measured distances $\delta_{i,j}^t$ and the actual distances based on location coordinates $d_{i,j}(X^t)$, which are calculated as,

$$d_{i,j}(X) = \|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\| = \sqrt{(\boldsymbol{x}_i^t - \boldsymbol{x}_j^t)^T (\boldsymbol{x}_i^t - \boldsymbol{x}_j^t)}. \quad (1)$$

Whenever a node's velocity changes, its non-zero acceleration is measured by its accelerometer. We incorporate this extra information in the latter part of the sum, representing the error between the measured acceleration $\boldsymbol{\alpha}_i^t$ and acceleration $\boldsymbol{a}_i^t$ which calculated from the coordinate path travelled as:

$$\boldsymbol{a}_i(X) = \left(\boldsymbol{x}_i^{t+1} - \boldsymbol{x}_i^t\right) - (\boldsymbol{x}_i^t - \boldsymbol{x}_i^{t-1}). \quad (2)$$

Our approach finds $\hat{X} = \text{argmin}_X S(X)$, in a distributed manner, to provide location estimates $\{\hat{\boldsymbol{x}}_{\boldsymbol{i}}^{\boldsymbol{t}}\}$.

Note that $S_i(X)$ is local to i^{th} node since it only depends on the measurements available at i^{th} node and positions of its neighbour nodes. Minimizing $S_i(X)$ with respect to $\{\boldsymbol{x}_i^t\}_t$ results in new position estimates for node i. Implementing our approach at each node constructs the backbone of this distributed method. We use majorization at node i to guarantee non-increasing local cost.

Our method is described in Algorithm 1 in [11]. The algorithm is iterative and must be given initial position estimates. Generally, each time the algorithm is run, it is initialized using the coordinates of positions from the previous round's estimates, $\boldsymbol{x}_i^t$ for $i = 0, \ldots, N-1$. The first time a neighbor j appears to node i, it must provide its own locations, $\{\boldsymbol{x}_j^t\}_{t=0,\ldots,N}$. Here, we use classical MDS to generate any coordinates $\{\boldsymbol{x}_j^t\}$ for which there are no prior round estimates.

B. Regression Based Collision Prediction Algorithm

Using the relative locations estimated from previous stage, we predict locations into the near future. Regression analysis is widely used for prediction and forecasting, as it reveals the causal relationships between a dependent variable and one or a collection of independent variables. We choose quadratic regression since trajectories are quadratic in constant acceleration, and polynomial regression generally works well for non-linear interpolation problems. In our case, we predict the future trajectory of the node, which has a non-linear relationship with time due to the dynamic nature of the motion. The output of the polynomial regression in such a scenario can also be interpreted as higher order kinematics. Given T data points $(t, \boldsymbol{x}_i^t)$, where the independent variable t is a time instance and $\boldsymbol{x}_i^t$ is the corresponding location of node i at times $t \in \{1, \ldots, T\}$, we fit a 2^{nd} degree polynomial to approximate node i's location for real-valued t,

$$\hat{\boldsymbol{x}}_i(t) = \boldsymbol{p}_2 + \boldsymbol{p}_1 t + \boldsymbol{p}_0 t^2, \quad (3)$$

where, $\boldsymbol{p}_0$, $\boldsymbol{p}_1$, and $\boldsymbol{p}_2$ are the polynomial coefficients,which we estimate using the least squares approximation giving the estimated coordinates $\hat{X}$. Using the coefficients, the algorithm extrapolates future relative locations of each node, thereby, giving future inter-object distances of each pair of nodes. We are interested in the *near future*, i.e., the time-frame that is

equal to or less than the reaction time of the node, which is application dependent. If the minimum inter-node distance threshold between two nodes is crossed within this near future, a collision is predicted.

IV. Experiments

A. Hardware

We conduct a series of experiments with mobile nodes to predict collisions between any pair of nodes. Each node follows the architecture as described in [12]. Each such node is attached to a iRobot Create that moves the node as programmed. Lastly, a Raspberry Pi3 processor is attached on top of each node, which both lets us program the node movement, and measures the acceleration of each node via a BN0055 IMU sensor [13].

B. Multi-node Ranging Protocol

Each node measures ranges between itself and all other nodes, and no anchors are present in the system. An efficient way to measure all $\binom{N}{2}$ ranges between the N nodes is to use the efficient multi-node ranging protocol in [6], which requires only N message exchanges per cycle to get all the ranges.

C. Setup

We set $N = 4$ floor nodes to move as depicted in Figure 1 in a 6m $\times$ 6m area. Each mobile node (top right in Figure 1) undergoes acceleration as detailed in Table I, constantly between its starting position and its stopping position in Test I and II. Test III has node 3 in motion at constant speed, and due to its motion in a circle, the magnitude of its acceleration is 0.125 m/s^2. We collect UWB ranges between every pair of nodes at a rate of 18 ranges per second. The acceleration of each node is measured via the IMU sensors and collected by their attached Raspberry Pi at a rate of 100 samples per second. We route the ranges and acceleration data collected by each node to a central processing unit for offline algorithm testing and result generation. Note that this offline implementation is just for convenience during tests; our distributed algorithm can be implemented in firmware at each node, and will be our future work. Each Raspberry Pi is NTP time synchronized, such that the timestamps for ranges and acceleration can be matched to produce 18 range-acceleration pairs per second. Lastly, to record the ground truth coordinates of each node during each experiment, we use a 16-camera OptiTrack motion capture system, which enables millimeter accuracy [14]. The results are explained in Section V.

	Test I	**Test II**	**Test III**
Stationary Nodes	1	2	3
Mobile Nodes	3	2	1
Acceleration (m/s^2)	0.125, 0.09, 0.06	0.125, .06	0.125

TABLE I
Node Setup in Three Tests

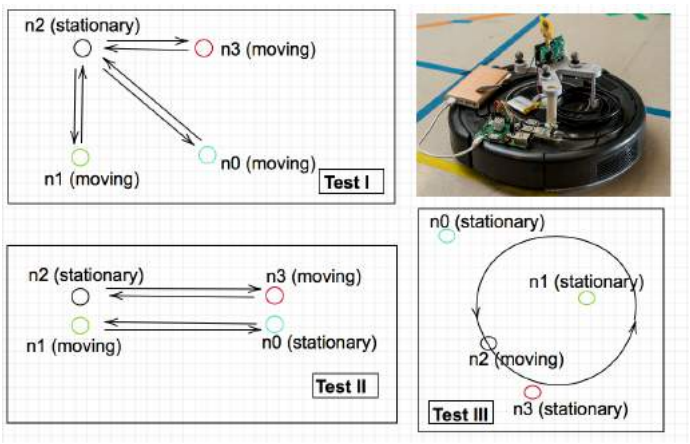


Fig. 1. Overview of the motions conducted by each node in three tests, and (Top Right) photo of hardware setup.

V. Results

A. Location Estimation

We first demonstrate the quality of location estimates generated from the ACED algorithm. For any of the 4 nodes, ACED estimates the trajectory that was followed over time. We use a window of $T = 20$ samples, thus each time the algorithm is run, we estimate $\{\boldsymbol{x}_i^t\}$ for $t = 1, \ldots 20$ and $i = 1, 2, 3, 4$. In a sliding window manner, ACED repeats by dropping the oldest time and adding one new time, and re-running the estimation for the next window of time. As a typical example, we plot the T location estimates for node 2 as '$\times$' and the ground truth locations as '$\circ$' in Figure 2. We also plot the location estimates from another state-of-the-art method, friend-based autonomous collision prediction and tracking (FACT) [6]. FACT assumes a constant velocity, and hence is unable to track the curved trajectory of node 2 from Test III at all. Furthermore, ACED is capable of predicting future positions based on (3), which are plotted for node 2 in the Figure 2 against the ground truth, where we define 'near future' as within 0.02 sec into the future.

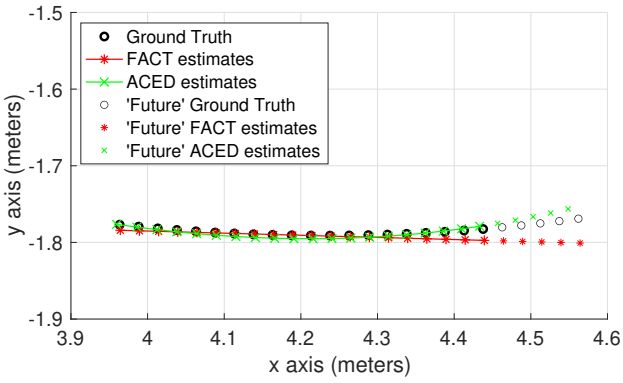


Fig. 2. Location estimates by ACED and FACT for the node moving with acceleration per window. ACED's predicted 'future' location estimates, are also plotted against ground truth.

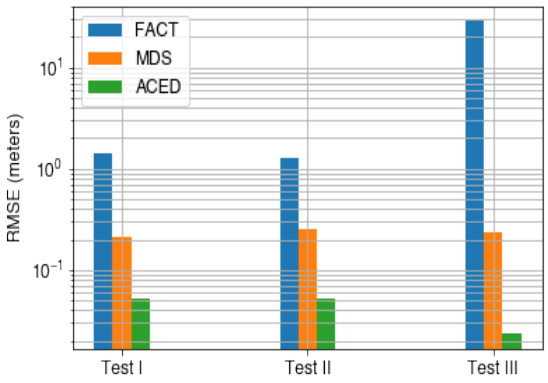


Fig. 3. Comparison of ACED vs. the modified MDS [5] and FACT [6] for each test, showing RMSE averaged across all nodes.

We use the following as our metric of error for the output of one window of any estimator:

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^{N} \left\| x_i^{T/2} - \hat{x}_i^{T/2} \right\|^2 \right]^{1/2}, \qquad (4)$$

and we report the average RMSE across all sliding windows across the entire test. Figure 3 plots this average RMSE of location estimated for all the nodes during each test by three methods, ACED, FACT[6] and modified MDS[5]. From [5], we used its centralized algorithm to find a globally optimum solution for relative position and velocity; however, it is known to perform sub-optimally in noise [5], [6].

Our results show that the FACT method of [6] diverges over time when motion is not linear. As described, each new run of the algorithm uses as initialization the trajectory estimates from the prior window. In Test III with a circular track for node 2, as FACT estimates a linear trajectory, its initialization from the prior window is poor. Over the course of Test III, its estimates at some point are unable to converge to the global optimum, after which it loses track of the coordinates and is unable to recover, leading to a very high average RMSE across Test III. ACED provides better tracking in dynamic motion, and doesn't have this convergence problem in our experiments. ACED also compares well to the centralized modified MDS method of [5], demonstrating a lower RMSE by 5-10x.

B. Interpolated Distance-based Collision Prediction

In order to avoid collision, a node must predict a collision before it happens. In ACED, if the distance $\hat{d}_{i,j}^t$ between nodes i and j in the near future t will fall below a threshold, d_{thd}, this counts as a future collision. Using the relative location estimates from ACED, we extrapolate pairwise distances into the near future. In this experiment, we define the near future as within $\tau = 0.02$ s.

We set the distance threshold d_{thd} to allow a trade-off between false alarms and missed detections. Letting r be the radius of one autonomous object, we would set $d_{thd} = 2r + \epsilon_d$ for some $\epsilon_d \geq 0$. By increasing ϵ_d, we would increase the probability of detection of a collision P_D while also increasing the probability of false alarm P_{FA}. A user could set the threshold based on the desired trade-off between the two. Figure 4 shows the ROC curve, i.e., the relationship between P_D and P_{FA}, compiled with data from across all three tests. ACED is able to provide higher P_D for the same P_{FA} when compared with FACT [6]. Note that even when FACT location estimates diverge, it manages to keep an accurate relative position and velocity for two nodes that are very close, and thus collision predictions are good. However, ACED cuts P_{FA} approximately by a factor of 2 for a constant P_D compared to FACT. Since ACED provides more accurate kinematics whenever nodes are accelerating, it can extrapolate complicated trajectories better, thus providing accurate future inter-node distances and kinematics to predict collisions. The 2^{nd} degree regression coefficients are able to extrapolate the future locations while taking each node's acceleration into account, a trait not achievable by FACT. We also test against the pairwise method of [15], which does not perform nearly as well as FACT or ACED.

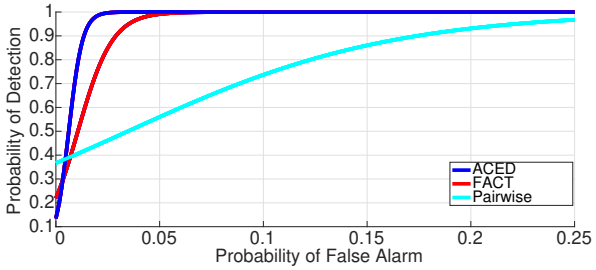


Fig. 4. ROC plot comparing ACED, FACT [6], and pairwise [15].

VI. Conclusion

This paper presents ACED, a new approach to estimate trajectories and predict collisions for systems involving mobile devices that are capable of measuring node acceleration and pairwise range measurements.The algorithm does not require a known-location infrastructure or a centralized computation. We test its performance in a network of four prototype mobile nodes mounted on ground robots in three tests. ACED predicts a node's trajectory with an order of magnitude lower RMSE, and collisions with a $> 2x$ lower false alarm probability, than three state-of-the-art infrastructure-free trajectory estimation and collision prediction methods.

Acknowledgment

This work is supported in part by the US National Science Foundation under Grant No. 1622741. We thank Alemayehu Solomon Abrar for sharing his code with us.

References

[1] D. Daneshvar, C. Nowinski, A. Mckee, and R. Cantu, "The epidemiology of sport-related concussion," *Clinics in sports medicine*, vol. 30, pp. 1–17, vii, 01 2011.

[2] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, "Localization from mere connectivity," in *Proc. 4th ACM Intl. Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '03, 2003, p. 201–212. [Online]. Available: https://doi.org/10.1145/778415.778439

[3] J. A. Costa, N. Patwari, and A. O. Hero III, "Distributed multidimensional scaling with adaptive weighting for node localization in sensor networks," *IEEE/ACM Transactions on Sensor Networks*, vol. 2, no. 1, pp. 39–64, Feb. 2006.

[4] B. Beck, R. Baxley, and J. Kim, "Real-time, anchor-free node tracking using ultrawideband range and odometry data," *Proceedings - IEEE International Conference on Ultra-Wideband*, pp. 286–291, 11 2014.

[5] R. T. Rajan, G. Leus, and A.-J. van der Veen, "Relative kinematics of an anchorless network," 2018.

[6] A. S. Abrar, A. Luong, G. Spencer, N. Genstein, N. Patwari, and M. Minor, "Collision prediction from uwb range measurements," *arXiv preprint arXiv:2010.04313*, 2020.

[7] P. Wei, L. Cagle, T. Reza, J. Ball, and J. Gafford, "Lidar and camera detection fusion in a real time industrial multi-sensor collision avoidance system," 2018.

[8] A. Viquerat, L. Blackhall, A. Reid, S. Sukkarieh, and G. Brooker, "Reactive collision avoidance for unmanned aerial vehicles using Doppler radar," in *Field and Service Robotics: Results of the 6th International Conference*, C. Laugier and R. Siegwart, Eds. Springer Berlin Heidelberg, 2008, pp. 245–254.

[9] R. Chellappa, Gang Qian, and Qinfen Zheng, "Vehicle detection and tracking using acoustic and video sensors," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 2004, pp. iii–793.

[10] I. Borg and P. Groenen, "Modern multidimensional scaling: Theory and applications, volume 2 of statistics in social science and public policy," 1997.

[11] A. Singh and N. Patwari, "Range-based collision prediction for dynamic motion," in *2021 IEEE 18th Annual Consumer Communications Networking Conference (CCNC)*, 2021, pp. 1–6.

[12] A. Luong, P. Hillyard, A. S. Abrar, C. Che, A. Rowe, T. Schmid, and N. Patwari, "A stitch in time and frequency synchronization saves bandwidth," in *ACM/IEEE Intl. Conference on Information Processing in Sensor Networks (IPSN 2018)*, April 2018, pp. 96–107.

[13] K. Townsend, *Adafruit BNO055 Absolute Orientation Sensor*. [Online]. Available: https://learn.adafruit.com/adafruit-bno055-absolute-orientation-sensor/overview

[14] NaturalPoint. Motion capture systems - optitrack webpage. [Online]. Available: optitrack.com

[15] A. S. Abrar, N. Patwari, and J. Decavel-Bueff, "Demo abstract: Collision prediction from pairwise ranging," in *19th ACM/IEEE Intl. Conference on Information Processing in Sensor Networks (IPSN 2020)*, April 2020.

Design and Simulation of an Autonomous Racecar: Perception, SLAM, Planning and Control

Sihao Wu, Zhengwei Yang, Xiaopo Xie, Yilong Wang, Xinliang Wang, Qi Wang, Bofan Wu, Hongjun Zhang, Hanning Zhang, Haochun Ma, Xuanliang Zhang and Haiying Lin

AERO Driverless Racing Team, Beihang University

ABSTRACT

Formula Student Driverless competition focuses on design, manufacture and racing of an autonomous racecar on closed track. This paper describes the hardware and software concept of the 2020-rebuilt AERO Driverless Racing Team. To achieve driverless goal, our autonomous racecar is modified based on the 2017 AERO electric racecar by adding the brake-by-wire and steer-by-wire. The driverless software system can be divided into perception, SLAM, planning and control. To speed up software development, we built our simulation model on the Vrep simulator containing vehicle model, autonomous sensors, TF relationship and racing track environment. This paper is a summary for past effort and a start of future development.

Index Terms— Autonomous Racecar, Software, Simulation

1. INTRODUCTION

Recent years have witnessed a rapid process of autonomous driving, among which the Formula Student Driverless competition (FSD) focuses on design, manufacture and racing of an autonomous racecar on closed track. All of the universities need to construct their own racecars under the rules set by the organizing committee of FSD [1] [2]. This paper describes the hardware and software concept of the 2020-rebuilt AERO Driverless Racing Team.

To achieve driverless goal, our autonomous racecar is modified based on the 2017 AERO electric racecar by adding the brake-by-wire and steer-by-wire. The specific hardware concept for autonomous driving contains GPS/IMU, LiDAR, monocular camera and central computing unit, as shown in Fig. 1.

Our work on the autonomous software system will be introduced on this paper as well. There are two pipelines on perception part, which are camera-based and LiDAR-based. For localization and mapping, we utilize ICP algorithm to match the purified points from the LiDAR data. Next, to obtain the midline of the racing track, we present a planning framework based on Delaunay triangulation and tree-search algorithm. Finally, we use pure-pursuit algorithm to achieve latitudinal control for racecar.

A highly emulated simulation has great potential for sub-system optimization, as well as system verification without labor and time waste. In our development process, we built our simulation model on the Vrep simulator and co-simulated based on ROS and Vrep.

Fig. 1: AERO Autonomous Racecar at the Formula Student China 2020. The location of LiDAR, GPS/IMU, monocular camera and central computing unit are marked in order as 1 to 4.

The contributions of this paper are:

- development of hardware concept and software system.
- software co-simulation based on ROS and Vrep.

The remainder of this paper is organized as follows. Section II presents state-of-the-art work of autonomous vehicle. In Section III, the hardware and simulation concept of our racecar will be introduced. Section IV describes the software concept and Section V presents our simulation results. The last part is our conclusion and future work.

2. RELATED WORK

The earlier autonomous racing can be traced back to the DARPAR competition, which asked the vehicles race in the desert without human assist [3]. Although the history of FSD is much shorter compared with DARPAR, there are still numerous studies about the autonomous racecar [8][17]. In the remaining paragraphs, related literatures for each sub-system of autonomous vehicle will be provided.

The camera-based detection is responsible to identify the color and preferably the location of the cones within the FSD competition. It has been verified that YOLO v3 [4][19] can adapt to the small-scale obstacle detection in real-time application. LiDAR detection is more redundancy compared to camera. AMZ utilized Euclidean clustering and classification based on geometric priors to detect the cones in [5]. Additionally, LiDAR-based SLAM is the most effective SLAM method nowadays [6], which uses LiDAR as the perception sensor. Zhang et al. [7] proposed LOAM that leads on the KITTI odometry benchmark recently. The

planning method is responsible for finding out a safety drivable route, pushing racecar to its limits. AMZ has proposed graph search path planning algorithm in [8]. The planning method used on our racecar is the same as the AMZ method. The control techniques generally can be divided into two types, pure-pursuit algorithm [9][16] and Model Predictive Control (MPC) [10][18].

3. FRAMEWORK OF AUTONOMOUS RACECAR

According to the FSD rules, the self-developed autonomous vehicle should complete the racing competition without track information. Therefore, we automated our electric vehicle on the hardware concept firstly. To speed up software iteration, we has modeled our autonomous racecar on Vrep simulator.

3.1. Hardware Concept

The hardware concept contains actuators, sensors and central computing unit, and the location of them is shown as Fig. 1. Due to the length of article, the details about central computing unit hasn't been shown.

A) Actuators

- Brake-by-wire

The brake-by-wire system contains servomotor, reducer gearbox and cable, as shown in Fig. 2(a). The cooperation between the servomotor and reducer gearbox can output torque about 61Nm and linearly control the brake pedal by cable. Although this is a quite simple layout, it is very effective for our brake-by-wire system.

(a) (b)

Fig. 2: (a) Brake-by-wire model. (b) Steer-by-wire model.

- Steer-by-wire

The main part of the steer-by-wire system is the steering servomotor connected to the steering gearbox by the electric-clutch (see Fig. 2(b)). The opening and closing of the clutch achieves mode switching between manual mode and driverless mode.

B) Sensors

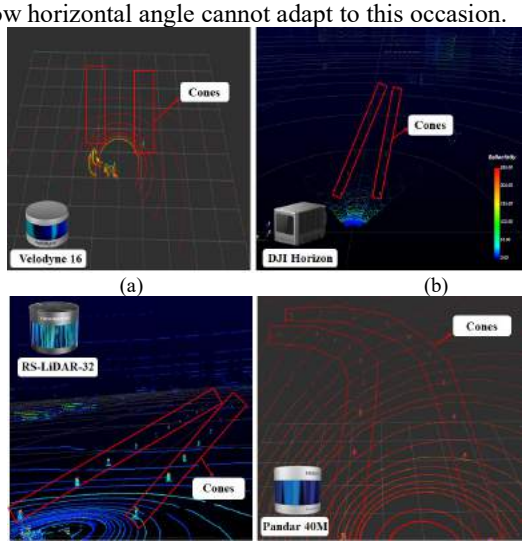
To automate our racecar for the competition, we equip monocular camera and LiDAR sensors, which are responsible for identifying object color and location respectively.

- Monocular camera

The main function of monocular camera is to identify color of cones. Since we use the LiDAR to detect the cone's distance, the stereo camera solution is given up. We chose solution of lens with focal length as 8mm and H×V as 58.4°×44.6°.

- LiDAR

The performance of LiDAR detection is greatly affected by various parameters. The forward-looking distance and vertical resolution will impact the farthest visibility distance. While the horizontal angle of view represents the width perception ability for the vehicle. Since there will be hair-pin circles on the racing track, the LiDAR with low horizontal angle cannot adapt to this occasion.



(a) (b)

(c) (d)

Fig.3: LiDAR test results about Velodyne 16 (a), DJI Horizon (b), RS-LiDAR-32 (c) and Pandar 40M (d).

Therefore, several LiDARs has been tested, consisting of Velodyne 16[11], DJI Horizon [12] RS-LiDAR-32[13] and Pandar 40M [14]. The results shown that Velodyne 16 can only detect the first two cones (see Fig. 3(a)), as it has just 16 channels leading to a low vertical resolution. Although DJI Horizon has a good performance on detection distance shown in Fig. 3(b), its horizontal angle is about 81.7°. From Fig. 3(c) and Fig. 3(d), we found that both RS-LiDAR-32 and Pandar 40M can meet our requirement. Considering stability, we chose Pandar 40M finally, which is sponsored by Hesai Company.

3.2. Simulation platform

The advantage of simulation is effectively shortening development time and reducing test cost. The most critical point is that the code updating can be performed while the vehicle is under construction. We developed the racecar on the Vrep simulator [15], covering from vehicle model, perception concept, GPS/IMU and TF relationship. The racing track has also been built to test. The overview of simulation platform is shown in Fig. 4.

As Pandar 40 model has not updated in Vrep, the Velodyne 64 with the similar performance is chosen. The vision sensor uses the common camera, setting the horizontal field of view as 60 degree and the forward-looking distance as 30 meters. Both in reality experiment and simulation test, TF relationships need to be established between each sensors, car body and map, which is a critical

part in localization and sensor fusion. In Vrep, the map link is as our initial link connected with the odomotry link followed by base_link (the center of car body). The rest of the sensors and other components mounted on the car are directly connected to the base_link. Therefore, the relative pose between each coordinate system can be obtained through the established TF relationship (see Fig. 5).

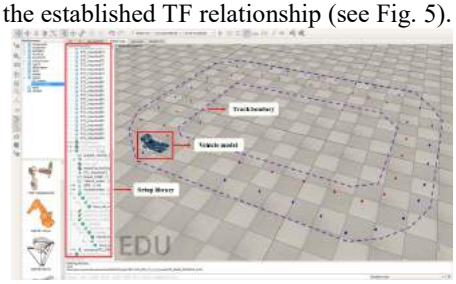


Fig. 4: Overview of autonomous racecar on Vrep simulator.

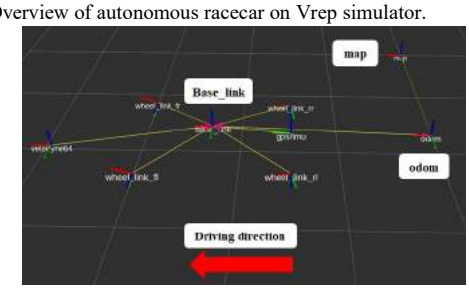


Fig. 5: TF relationship in Rviz visualization.

4. SOFTWARE METHOD

Generally, the track map should be computed in the first lap and later race alongside the previously mapped track. To this end, our software is divided into perception, planning and control, and the framework is shown in Fig. 6.

4.1. Perception

There are two pipelines in perception sub-model, which cannot only get the color information with camera but also obtain accurate location of cone with LiDAR.

A) Camera-based

Due to the high speed during racing, the real-time and one-step object detection algorithm, YOLO v3, is used to detect the cone. In the execution of YOLO v3, the classification-based CNN networks, as the backbone, are to determine the category of the object with given bounding box. The network has the characteristics of low computational cost and fast running speed while ensuring strong feature extraction strength. Before training, we use the KMeans method to cluster the bounding box in the dataset and get nine prior anchors that are employed to constrain the range of predicted objects. Then in the training process, the images are resized to 416*416 and the Adam optimizer with a learning rate of 1e-3 is used. To train the network, we collected 485 images in Vrep. Among them, 357 images are training/validation sets, and 128 images are test sets. We train the model for 50 epochs to make it converge.

B）LiDAR-based

To identify the location of the cone, the LiDAR detection algorithm has five steps (see Fig. 7). First, the interesting area is filtered out by the straight through filter. Then the ground segmentation is obtained by RANSAC. We use the data structure, KD-Tree, to preprocess the above ground points. Followed is the Euclidean Clustering and the projecting optimization. More specifically, the cluster tolerance is set as 0.5 and the cluster size is from 10 to 1000. While the CropHull method is utilized from PCL to restore points deleted by mistake during the ground segmentation. Finally, the information of point cloud and cone will be send to SLAM and planning respectively. In LiDAR-SLAM, the Point-to-Plane type, ICP method, is used to estimate the pose transformation of the filtered point cloud between adjacent frames. The rotation and translation matrix of the point cloud are obtained, furthermore getting the matching map and the estimated vehicle pose.

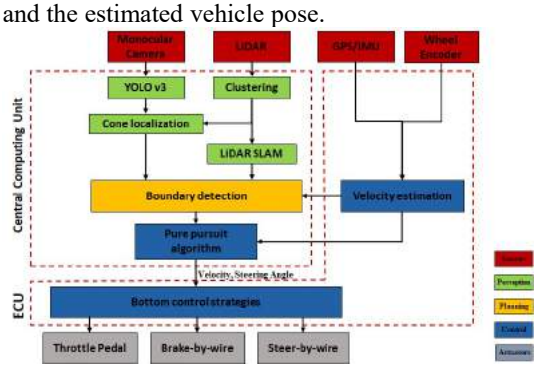


Fig. 6: Software framework indicating the relationship between sub-models of our driverless racecar.

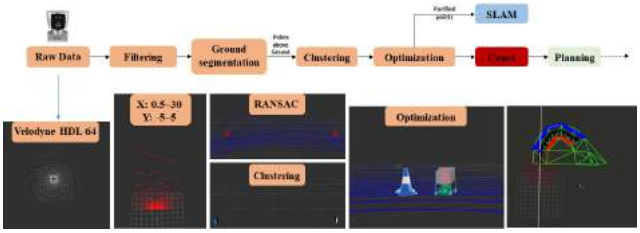


Fig. 7: The flow chart of LiDAR-based perception algorithm.

4.2. Path Planning

The planning problem can be transferred into a graph search problem, which contains three steps. The first step is to obtain the midpoint of the corresponding cone line from perception results, using the Delaunay triangulation method [20], as shown is Fig. 8(b). There are different routes combined with calculated midpoints shown as Fig. 8(c). Therefore, the tree-search algorithm is chosen, where node with the lowest cost function among all nodes is selected. The cost function is given by,

$$cost_{node}=cost_{color}\cdot w_c+\frac{cost_{theta}}{2\times M\_PI}\cdot w_t+cost_{width}\cdot w_w+cost_{distance}\cdot w_d \quad (1)$$

where $cost_{node}$, $cost_{color}$, $cost_{theta}$, $cost_{width}$ and $cost_{distance}$ respectively represent total cost of the node, penalty of the cone color, the penalty of connection angel, the penalty of

the connection length, and the penalty of the distance between nodes. w_c, w_t, w_w and w_d are preset parameters and represent the importance of the four types of punishment respectively.

Finally, the tree branch with the lowest penalty value is selected as the optimal path output based on filtered points (see Fig. 8(d)). The time complexity of the tree search algorithm is $O(n^m)$, where n, m respectively represent the depth and width of the tree search, set as 5 and 3.

4.3. Control

Obtaining the results of planning, we adopt the pure pursuit method to follow the path. By controlling the front wheel steering angle δ, the vehicle can drive along a circular arc passing through the preview point.

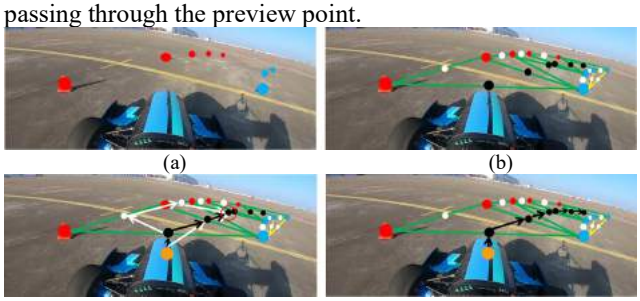


Fig. 8: (a) Results from Perception part. (b) Results of Delaunay triangulation. (c) Combination of calculated midpoints. (d) The selected path with the lowest penalty value.

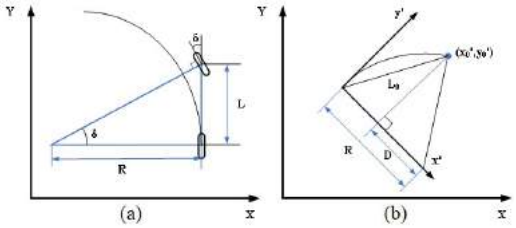


Fig. 9: (a) Kinematics model of two-wheel bicycle. (b) Geometric analysis for pure pursuit algorithm.

The Fig. 9(a) demonstrates the simplified bicycle model for vehicle kinematic model. The relationship between the front wheel steering angle δ, the wheelbase L and steering radius R is obtained as:

$$\tan(\delta)=\frac{L}{R} \tag{2}$$

In pure pursuit model (see Fig. 9 (b)), the vehicle coordinate system is $x'-y'$. (x_0', y_0') is the preview point on vehicle coordinate system and L_0 is the distance between rear axle and the preview point. The geometric relationship is described in equation (3). After determining R, the front wheel steering angle δ can be obtained, and then an instruction is issued to control the vehicle to follow the preview point.

$$R = \frac{L_0^{\,2}}{2x_0'} \tag{3}$$

5. SIMULATION RESULTS

In the following section are the results based on ROS and Vrep co-simulation platform.

TABLE I: Test results of YOLO v3

Image Size	FPS	AP50 $_{Red}$	AP50 $_{Yellow}$	AP50 $_{Blue}$	mAP50 $_{Total}$
416*416	56	0.967	0.993	0.968	0.976

Fig. 10 shown the outcome of camera-based detection. The Average Precision (AP) and mean Average Precision (mAP) are evaluation index of algorithm accuracy. The Frames per Second (FPS) is the evaluation index of real-time nature. Table I demonstrate that the mAP_{50} Total is about 0.976, which can meet our requirement. Fig. 11 is the result of LiDAR detection and SLAM. The red and blue cylinders represent different cones. Moreover, the red line is the estimated vehicle track line and the green points cloud is the built map, which are calculated by ICP-SLAM. Fig. 12 demonstrates the visualization of planning and control. As for planning, the green triangles are the outcome of Delaunay triangulation method, the blue and red thick lines are the estimated boundary. Meanwhile, the green and blue spheres represent current preview point and next preview point respectively. Moreover, the white curve is the track curve for our racecar.

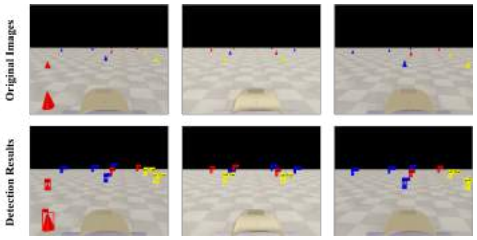


Fig. 10: Simulation results of Camera.

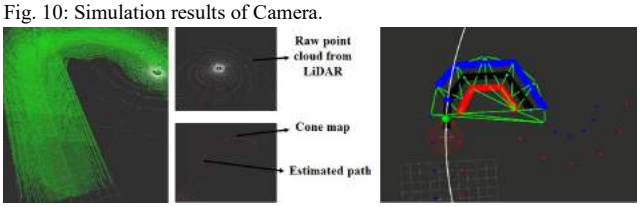


Fig. 11: Simulation results of LiDAR. Fig. 12: Visualization of planning and control result.

6. CONCLUSION

In this paper, the AERO driverless racecar was presented on its hardware and software concept. The co-simulation platform based on Vrep and ROS was developed to accelerate the development process, which is the first attempt in the field of autonomous racecar. The completed autonomous framework was constructed for 2020 competition. There are still several parts with the potential to be improved, which are state estimation, localization fusion and perception fusion. Meanwhile, we are building our new driverless racecar now. This paper is a summary for our past effort and a start of our future development.

ACKNOWLEDGEMENT

The authors thank the AERO driverless team for their hard work and contribution, and the support from the AERO electric team. We also appreciate the greatly support from the Transportation College of Beihang University. Finally, we really admire the contribution from BIT driverless team and AMZ driverless team to the FSD competition.

REFERENCE

[1] FSC competition handbook 2020. http://www.formulastudent.com.cn/
[2] FSG competition handbook 2020. https://www.formulastudent.de/fileadmin/user_upload/all/2020/rules/FS-Rules_2020_V1.0.pdf
[3] M. B. K. Iagnemma, “Special issue on the darpa grand challenge, part 2,” Journal of Field Robotics, vol. 23, pp. 661–692, September 2006.
[4] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” 2018.
[5] Design of an Autonomous Racecar: Perception, State Estimation and System Integration. Miguel I. Valls*, Hubertus F.C. Hendrikx*, Victor J.F. Reijgwart*, Fabio V. Meier*, Inkyu Sa, Renaud Dubé, Abel Gawel, Mathias Bürki, and Roland Siegwart.
[6] https://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping
[7] J. Zhang and S. Singh, "LOAM: Lidar odometry and mapping in real-time", Robotics: Science and Systems Conference (RSS), July 2014.
[8] J. Kabzan et al., "AMZ driverless: The full autonomous racing system," J. Field Robot, Aug. 2020.
[9] R. Craig Coulter, "Implementation of the Pure Pursuit Path Tracking Algorithm," Jan. 1992.
[10] Borrelli F, Falcone P, Keviczky T, et al., "MPC-based approach to active steering for autonomous vehicle systems," International Journal of Vehicle Autonomous Systems, pp. 265–291, Jan. 2005.
[11] https://velodynelidar.com/products/puck/
[12] https://store.dji.com/sg/product/livox-horizon-lidar?vid=89181
[13] https://www.robosense.ai/rslidar/RS-LiDAR-32
[14] https://www.hesaitech.com/en/Pandar40M
[15] CoppeliaSim User Manual, 2020. Available: https://www.coppeliarobotics.com/helpFiles/
[16] Jarrod M. Snider, "Automatic Steering Methods for Autonomous Automobile Path Tracking," Robotics Institute, Carnegie Mellon University, Feb. 2009.
[17] J. Ni, J. Hu and C. Xiang, "Robust Path Following Control at Driving/Handling Limits of an Autonomous Electric Racecar," in IEEE Transactions on Vehicular Technology, vol. 68, no. 6, pp. 5518-5526, June 2019.
[18] José L. Vázquez, et al., "Optimization-Based Hierarchical Motion Planning for Autonomous Racing," arXiv e-prints, Mar 2020.
[19] Redmon, Joseph , et al., "You Only Look Once: Unified, Real-Time Object Detection," Computer Vision & Pattern Recognition IEEE, pp. 779–788, Dec. 2016.
[20] B. Delaunay, "Sur la sphère vide," Science USSR VII:Class. Sci. Mat. Nat, pp. 793–800, 1934.

GRAPH-BASED MOTION PLANNING FOR AUTOMATED VEHICLES USING MULTI-MODEL BRANCHING AND ADMISSIBLE HEURISTICS

Oliver Speidel, Jona Ruof, and Klaus Dietmayer

Institute of Measurement, Control and Microtechnology
Ulm University, 89081 Ulm, Germany
firstname.lastname@uni-ulm.de

ABSTRACT

Automated driving in urban scenarios requires efficient planning algorithms able to handle complex situations in real-time. A popular approach is to use graph-based planning methods in order to obtain a rough trajectory which is subsequently optimized. A key aspect is the generation of trajectories implementing comfortable and safe behavior already during graph-search, while keeping computation times low. To capture this aspect, on the one hand, a branching strategy is presented in this work that leads to better performance in terms of quality of resulting trajectories and runtime. On the other hand, admissible heuristics are shown which guide the graph-search efficiently, where the solution remains optimal.

***Index Terms*—** Motion Planning, Trajectory Planning, Decision-making, Automated Vehicles, Autonomous Driving

1. INTRODUCTION

Over the last years, intensive research has been carried out in the field of autonomous driving [1, 2, 3]. Thereby, motion planning is a crucial requirement and one of the most challenging aspects for automated vehicles. As early as 2007, impressive automated systems for complex urban scenarios with interacting vehicles were presented as part of the well-known Urban Challenge initiated by the Defense Advanced Research Projects Agency (DARPA) [1]. In 2013, the Mercedes S-Class Bertha was able to drive fully autonomously more than $100\,\mathrm{km}$ from Mannheim to Pforzheim in Germany [2, 4]. A popular architecture for the motion planning system follows the idea that a behavior planning module decides for a strategic maneuver option, which is passed to a trajectory planning module where a feasible trajectory is calculated. A practicable approach for behavior planning is to generate a maneuver option rule-based using heuristics. However, this limits the capabilities for foresighted motion planning in complex environments [5]. For this reason, more foresighted but still efficient behavior and trajectory planning systems are widely investigated. A popular concept for graph-based behavior planning is shown in [6]. A speed profile along a given reference path is obtained by using graph-search methods. The idea is to extract a rough behavior trajectory over a planning horizon $t_{\mathrm{hor}} \approx 10\,\mathrm{s}$ in order to enable foresighted behavior planning. There exist several approaches which extend this concept of behavior planning for, e.g., short term lateral motion [7], merging behavior at highways [8] or courteous behavior at intersections [9]. Similar approaches formulate Partially Observable Decision Processes (POMDPs) which allow a sophisticated consideration of uncertainties [10, 11]. A key aspect of these approaches is still the trade-off between runtime and trajectory quality. Further concepts integrate domain knowledge into the search-graph by using models as the Intelligent Driver Model (IDM) [12] or the Pure Pursuit Controller [13] to improve trajectory quality [10, 14, 15]. However, a detailed analysis how to prune the available ego actions efficiently using driver models with minor reduction of trajectory quality is not regarded.

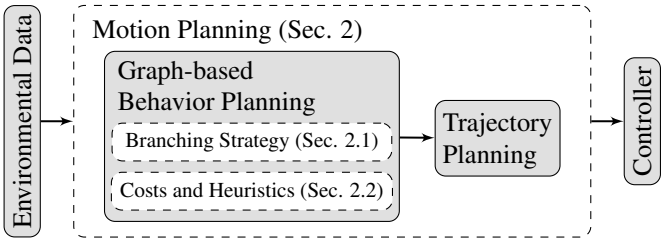


Fig. 1: System Overview

Based on the previous discussion, in this work, a motion planning framework is developed enabling foresighted and courteous behavior using graph-search methods extending our concept presented in [9]. The main idea is to integrate domain knowledge provided by different control and driver models into the existing framework to improve trajectory quality and efficiency. Therefore, a branching strategy is presented as well as admissible heuristics which are even applicable in interactive urban scenarios. As a result, we are able to improve the performance of driven trajectories and runtimes compared to related work.

2. METHODOLOGY

The concept follows the modular architecture of behavior and trajectory planning as shown in Figure 1. Preceding modules provide environmental data including map data as well

as state estimations of other vehicles with according predictions. Thereby, a set of predicted trajectories for each other vehicle with corresponding uncertainties is received.

The goal of the graph-based behavior planning is to obtain a rough behavior trajectory. In general, planning is done relative to the center line of the current road lane. Therefore, a node in the graph is represented by a state vector

$$\boldsymbol{x}_k = [s_k, d_k,\ \theta_k, \kappa_k, v_k, a_k]^\mathsf{T}, \quad k \in [0, \dots, T], \quad (1)$$

where s is the longitudinal position along the lane, d the lateral distance to the lane, θ the orientation, κ the curvature, v the velocity, a the acceleration and k the corresponding time step. The end of the planning horizon t_{hor} is denoted by the index T. The lane relative position $[s, d]$ can be transformed to the classic representation $[x, y]$ in Cartesian coordinates and vice versa. For further details, the reader is referred to [16].

The expansion of a node in the graph, i.e. the generation of possible subsequent states, is done using different models which will be presented in Section 2.1. The time discretization of subsequent nodes is $\Delta t = 1\,\text{s}$. Beginning from the root node, i.e. the current state, a graph is generated up to the planning horizon $t_{\text{hor}} \approx 10s$. The graph structure is exemplary shown in Figure 2. The optimal behavior trajectory is extracted using the A*-search algorithm, where the search is guided by the admissible heuristic functions shown in Section 2.2. The generation of the resulting trajectory in the trajectory planning module is based on the approach presented in [9]. The general idea is to use polynomials in order to interpolate between the behavior trajectory states. In contrast to our previous work, in this work we also regard lateral optimization. In the end, the resulting trajectory is passed to the controller which generates the input for the actuators.

2.1. Branching Strategy

In general, the branching strategy determines which actions to apply at each node to generate subsequent nodes and therefore, defines the graph structure. The idea is to combine the advantages of pre-defined acceleration actions that are used in [9] and model-based actions which generate preferable behavior for different scenarios, inspired by [15]. In order to omit the expansion of all actions, only a subset of actions is expanded at each node to reduce the graph size. In the following, this process of choosing which action to expand at which node is also referred to as *action selection*. In general, the ideas of [15] are extended by additional control models as well as more sophisticated action selection strategies. Further, in contrast to [15], the behavior planning is embedded into a holistic framework generating comfortable trajectories. In addition, the solution of behavior planning guarantees the existence of a feasible solution in the trajectory planning module, as all kinematic and collision constraints are considered during the expansion of an action.

In general, longitudinal actions $\mathcal{A}_{\text{lon}}$ and lateral actions $\mathcal{A}_{\text{lat}}$ can be distinguished, where the resulting action set is

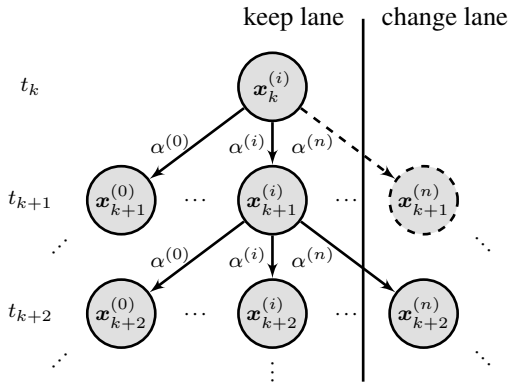


Fig. 2: Exemplary part of the graph utilized for behavior planning. Nodes represent states connected by edges associated with actions contained in $\mathcal{A} = \{\alpha^{(0)}, \dots, \alpha^{(n)}\}$. The concept of MOBIL-based action selection is demonstrated for a possible lane change maneuver, which is not expanded originating from $\boldsymbol{x}_k^{(i)}$ indicated by dashed lines. However, a possible lane change is regarded originating from $\boldsymbol{x}_{k+1}^{(i)}$.

$\mathcal{A} = \mathcal{A}_{\text{lon}} \times \mathcal{A}_{\text{lat}} = \{\alpha^{(0)}, \dots, \alpha^{(n)}\}$. Hereafter, the different actions and corresponding action selection strategies are presented.

Longitudinal: For the generation of longitudinal actions, acceleration α^a_{lon} and velocity targets α^v_{lon} are distinguished.

First, the acceleration targets are discussed, which are defined as $\alpha^a_{\text{lon}} \in \{-2, -1, 0, 1, 2, a^{\text{idm}}\}\,\text{m}\,\text{s}^{-2}$. These consist of pre-defined accelerations and the acceleration according to the IDM a^{idm}. The a^{idm} target is only expanded in car-following scenarios, where it is able to model comfortable and human-like following behavior. In order to omit expansions of actions similar to a^{idm}, in car-following scenarios only pre-defined acceleration targets with $|\alpha^a_{\text{lon}} - a^{\text{idm}}| > 0.5\,\text{m}\,\text{s}^{-2}$ are expanded. Further, we require $|\alpha^a_{\text{lon}} - a_k| \leq 1.9\text{m}\,\text{s}^{-2}$. The thresholds are determined based on simulation experiments and result in a reasonable trade-off between trajectory quality and runtime. The longitudinal state transition model for acceleration targets is defined by

$$\begin{bmatrix} s_{k+1} \\ v_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_k \\ v_k \\ a_k \end{bmatrix} + \begin{bmatrix} \frac{1}{6}\Delta t^3 \\ \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix} \dot{a}_k\,, \quad (2)$$

where $\dot{a}_k = (\alpha^a_{\text{lon}} - a_k)/\Delta t$ and, as a result, $a_{k+1} = \alpha^a_{\text{lon}}$.

The velocity targets $\alpha^v_{\text{lon}} \in \{v_0, v_d\}$ are defined by the reference velocity v_d and stillstand velocity $v_0 = 0\,\text{m}\,\text{s}^{-1}$, where the acceleration of the resulting subsequent states is constrained to $0\,\text{m}\,\text{s}^{-2}$. Therefore it applies $v_{k+1} = \alpha^v_{\text{lon}}$ and $a_{k+1} = 0\,\text{m}\,\text{s}^{-2}$. These actions are expanded if the target velocity is reachable within Δt. As state transition model, the concept of C1-continuous time optimal trajectories summarized in [17] is employed. Thus, comfort is ensured by restricted and continuous jerk. In general, C1-continuous time optimal trajectories also allow emergency maneuvers at kine-

matic limits. However, in this work, we limit our scope to non-safety critical scenarios. For further details, the reader is referred to [17].

Lateral: The lateral action set is given by the different road lanes which can be targeted to drive on. Therefore, $\mathcal{A}_{\text{lat}} = \{r_\text{l}, r_\text{c}, r_\text{r}\}$ where r_l represents the lane to the left, r_c the current lane and r_r the lane to the right. Using r_c the motion can be modeled purely longitudinal along the center line of the current lane. In order to perform a lane change to r_l or r_r, the Pure Pursuit Controller [13] is employed as lateral transition model and, consequently, the vehicle state is regarded in Cartesian Coordinates. The steering behavior defined by the Pure Pursuit Controller is combined with different acceleration targets for longitudinal behavior. This allows to restrict κ, $\dot{\kappa}$ and the absolute acceleration a_abs already during behavior planning which ensures feasible solutions in the trajectory planning module. In general, one lane change is allowed during $0.5t_\text{hor}$. The context in which a lane change is explored, is defined by the MOBIL model [18], which is known to generate human-like decision-making for lane change behavior [19]. Thereby, it is estimated if a lane change is favorable for the combined costs of all involved vehicles.

2.2. Cost and Heuristic Functions

Now as the graph structure is described, costs have to be defined in order to be able to apply graph-search methods. The costs attributed to a node or respectively a state are defined by

$$J = w_\text{f} j_\text{f} + w_\text{c} j_\text{c} + w_v j_v + w_a j_a + w_{\dot{a}} j_{\dot{a}} + w_\text{lc} j_\text{lc}\,, \quad (3)$$

where j_f represents costs for the spatio-temporal distance to the vehicle in front, j_c are courtesy costs that arise if the ego vehicle pulls out or drives in front of another vehicle, where possible interactions with the ego vehicle are considered. For further details, the interested reader is referred to [9, 20]. The difference to the desired velocity is regarded by j_v and the comfort is optimized by costs j_a and $j_{\dot{a}}$ for large absolute values of a and $\dot{a}$. Further, costs j_lc arise for lane changes. The single cost terms can be weighted with the according cost weighting w. In order to generate courteous and safe behavior, a set of predicted trajectories for each other vehicle is considered, where the corresponding uncertainties are incorporated by the single cost terms.

To further improve the runtime, admissible heuristics are developed which can be calculated online. The idea is to use a linear combination of heuristic terms rather than model directly one overall heuristic. If the heuristics for the single cost terms are admissible, the linear combination remains admissible [21]. Therefore, the heuristic is given with

$$h_\text{all} = \sum_{i=k+1}^{T} w_\text{f} j_{\text{f},i}^\text{min} + w_\text{c} j_{\text{c},i}^\text{min} + w_v j_{v,i}^\text{min} + w_a j_{a,i}^\text{min} + w_{\dot{a}} j_{\dot{a},i}^\text{min}\,, \quad (4)$$

where $j_{(\cdot),i}^\text{min}$ represent the minimal costs for the corresponding cost term that arise at time i originating from the currently regarded node $\boldsymbol{x}_k$. In the following, the calculation of the single minimal costs terms is explained. The terms $j_{\dot{a},i}^\text{min}$ and $j_{a,i}^\text{min}$ are determined by the minimal necessary jerk and acceleration to avoid a collision with the vehicle in front. The term $j_{\text{f},i}^\text{min}$ can be estimated by calculating the maximum possible distance to the vehicle in front at i. The same applies for $j_{\text{c},i}^\text{min}$, where the maximum possible distance to the vehicle behind is calculated. The minimal arising velocity costs $j_{v,i}^\text{min}$ are given if the ego vehicle accelerates with maximum acceleration to the desired speed. Even though the single minimal costs terms result in low estimated heuristic costs, the evaluation shows that the combination of all heuristics leads to a significant reduction of calculation times, while the solution remains optimal.

3. EVALUATION

The evaluation is done using real-world map data contained in a high-precision digital map of Ulm (Germany) [3]. The concept is implemented in C++ using the A*-search algorithm of the DOSL library [22]. Runtimes are obtained using a Intel XEON E5-1660 v4 CPU with 3.2 GHz utilizing a single thread. Other vehicles are simulated with random acceleration uniformly distributed between $[-1\,\text{m}\,\text{s}^{-2}, 1\,\text{m}\,\text{s}^{-2}]$ in each time step. In general, for each of the evaluations about 250 scenarios were analyzed, including lane change, highway on-ramp, roundabout and intersection scenarios.

3.1. Branching and Heuristic Functions

At first, the action selection strategy is investigated. The corresponding findings are summarized in Table 1. In order to measure the comfort of resulting trajectories, both average squared acceleration $\varnothing a^2$ and jerk $\varnothing \dot{a}^2$ are regarded, as they are also incorporated into the cost function during trajectory planning. The results show that our action selection strategy only has minor influence on the quality of resulting trajectories, while the runtime for behavior planning is reduced by 90% compared to a more passive action selection strategy similar to [15]. For both strategies, the maximum allowed acceleration difference and the number of lane changes is restricted as presented in Section 2.1. The results also emphasize that the MOBIL model yields well suited decision-making for lane changes when integrated into the graph-based framework. Consequently, the proposed action selection strategy enables the usage of the extended action set for real-time applications.

Further, the heuristic functions as well as the overall branching strategy were evaluated, where Table 2 shows corresponding results. As baseline, the branching strategy of the concepts [9, 6] is used. Because these do not regard lane changes, the proposed branching strategy is implemented for lateral actions. It is shown that comfort is much higher for the proposed approach, as $\varnothing a^2$ is slightly reduced and $\varnothing \dot{a}^2$ is nearly 25% lower. These results emphasize the idea that domain knowledge of driver and control models can be effectively used in order to improve comfort, in the trade-off

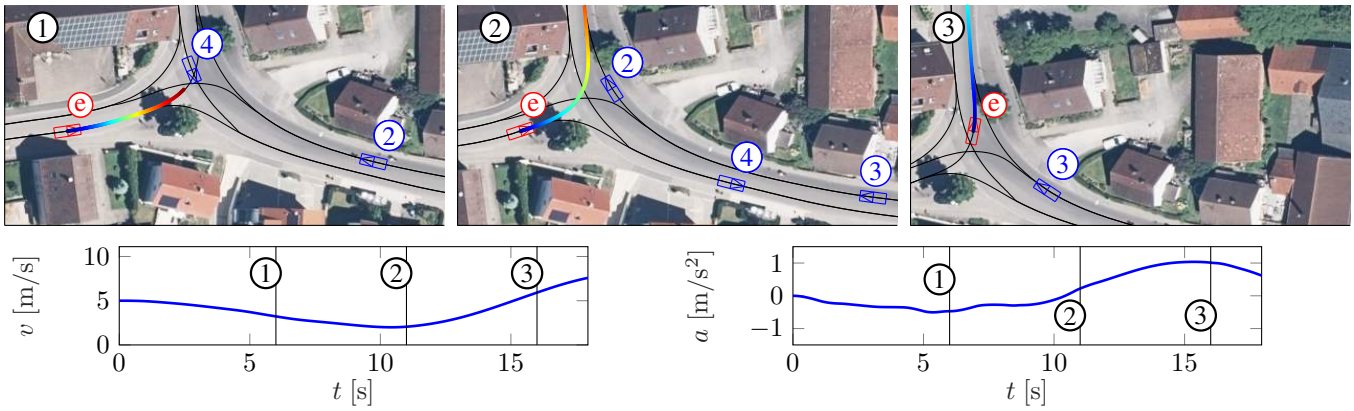


Fig. 3: Evaluation scenario. The ego vehicle performs a left turn, while maintaining comfortable behavior and courtesy towards other vehicles. In the first row, a top view of the scene is depicted for three different points in time. The planned trajectory is represented by colored dots where red presents the planned position at T. The ego vehicle is red and other vehicles are blue. The center lines of the road lanes are depicted in black. The second row shows the velocity and acceleration of the driven trajectory. Note that the driven trajectory is not the behavior trajectory but the result of the complete motion planning framework.

Table 1: Evaluation of the action selection strategy, including strategies for car-following and lane changing using the MOBIL model. The proposed action selection strategy is compared with a more passive one, i.e. less restrictive, similar to [15]. Mean and standard deviation (stdev) of runtimes (RT) for behavior planning as well as average squared jerk and acceleration for driven trajectories are shown. No heuristic functions (h_0) are used for the comparison.

mean ± stdev	proposed (h_0)	passive
∅ RT [ms]	**18.1 ± 4.9**	189.6 ± 168.4
max RT [ms]	**158.3**	1588.7
$\varnothing \dot{a}^2$ [(m/s^3)^2]	0.056 ± 0.023	0.056 ± 0.024
$\varnothing a^2$ [(m/s^2)^2]	0.39 ± 0.21	0.39 ± 0.21

against longer runtimes. Further, it is demonstrated that the average runtime for behavior planning can be reduced by about 17% using the proposed heuristic functions. The maximum runtime is even improved by about 60%, from 158.3 ms to 62.7 ms, while the solution of behavior planning remains optimal. Slight divergences of driven trajectories occur due to numerical issues. As a result, compared to [9, 6] the trajectory quality can be improved while the mean runtime is only slighlty increased. The maximum runtime can be even reduced by 39%, which allows for much faster replanning.

3.2. Motion Planning Framework

In order to give an insight to the overall performance and the resulting trajectories of the motion planning framework, an exemplary scenario is depicted in Figure 3. In general, an urban left turn scenario is regarded without right-of-way. At time ①, the ego vehicle ⓔ slowly approaches the intersection, while vehicle ④ crosses it. Afterwards, vehicle ② and ③ approaching from the right have to be considered. Taking the turn in front of vehicle ② would cause to much courtesy costs, thus the ego vehicle merges between vehicle ② and ③ at time ③. In addition, it is worth noting that during the analysis of all 250 scenarios, the maximum measured overall calculation time for motion planning was 89.33 ms using the presented approach. Further, non of the scenarios led to any collisions despite random behavior of other traffic participants. This emphasizes the capability of the framework to handle complex urban scenarios. Further evaluations and exemplary scenarios are presented in our video.[1]

Table 2: Evaluation of the proposed behavior planning module. The evaluation scenarios were investigated using the proposed branching strategy and the branching strategy implemented in [6, 9]. Further, results are shown using the proposed heuristic functions (h_{all}) and without the usage of heuristic functions (h_0), where resulting trajectories slightly deviate due to numerical issues.

mean ± stdev	proposed (h_0)	proposed (h_{all})	[6, 9]
∅ RT [ms]	18.1 ± 4.9	15.0 ± 2.9	**14.8 ± 3.0**
max RT [ms]	158.3	**62.7**	103.0
$\varnothing \dot{a}^2$ [(m/s^3)^2]	**0.056 ± 0.023**	**0.055 ± 0.022**	0.073 ± 0.025
$\varnothing a^2$ [(m/s^2)^2]	**0.39 ± 0.21**	**0.39 ± 0.21**	0.40 ± 0.22

4. CONCLUSION

In this work, we presented a motion planning framework for autonomous vehicles in urban environments utilizing graph-search methods. The proposed branching strategy and admissible heuristic functions yield trajectories attributed with lower costs, while the runtime is reduced significantly compared to related work. The implementation of the concept on the research vehicle of Ulm University and according validations in real-world public traffic is part of our future work.

[1] https://youtu.be/-VwqrXCJuP4

5. REFERENCES

[1] C. Urmson *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, 2008.

[2] J. Ziegler *et al.*, "Making bertha drive—an autonomous journey on a historic route," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 8–20, 2014.

[3] F. Kunz *et al.*, "Autonomous driving at ulm university: A modular, robust, and sensor-independent fusion approach," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, June 2015, pp. 666–673.

[4] J. Ziegler, P. Bender, T. Dang, and C. Stiller, "Trajectory planning for bertha 2014; a local, continuous method," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, June 2014, pp. 450–457.

[5] J. Ziegler, "Optimale bahn-und trajektorienplanung für automobile," 2015.

[6] C. Hubmann, M. Aeberhard, and C. Stiller, "A generic driving strategy for urban environments," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2016, pp. 1010–1016.

[7] W. Zhan, J. Chen, C. Chan, C. Liu, and M. Tomizuka, "Spatially-partitioned environmental representation and planning architecture for on-road autonomous driving," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 632–639.

[8] E. Ward and J. Folkesson, "Towards risk minimizing trajectory planning in on-road scenarios," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, June 2018, pp. 490–497.

[9] O. Speidel, M. Graf, T. Phan-Huu, and K. Dietmayer, "Towards courteous behavior and trajectory planning for automated driving," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Oct 2019, pp. 3142–3148.

[10] L. Zhang, W. Ding, J. Chen, and S. Shen, "Efficient uncertainty-aware decision-making for automated driving using guided branching," *arXiv preprint arXiv:2003.02746*, 2020.

[11] C. Hubmann, J. Schulz, M. Becker, D. Althoff, and C. Stiller, "Automated driving in uncertain environments: Planning with interaction and uncertain maneuver prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 5–17, March 2018.

[12] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical review E*, vol. 62, no. 2, pp. 1805, 2000.

[13] R. C. Coulter, "Implementation of the pure pursuit path tracking algorithm," Tech. Rep., Carnegie-Mellon UNIV Pittsburgh PA Robotics INST, 1992.

[14] A. G. Cunningham, E. Galceran, R. M. Eustice, and E. Olson, "Mpdm: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 1670–1677.

[15] D. Lenz, T. Kessler, and A. Knoll, "Tactical cooperative planning for autonomous highway driving using monte-carlo tree search," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 447–453.

[16] M. Werling, J. Ziegler, S. Kammel, and S. Thrun, "Optimal trajectory generation for dynamic street scenarios in a frenet frame," in *2010 IEEE International Conference on Robotics and Automation*, May 2010, pp. 987–993.

[17] K. L. Knierim and O. Sawodny, "Real-time trajectory generation for three-times continuous trajectories," *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1462–1467, 2012.

[18] A. Kestinge, M. Treiber, and D. Helbing, "General lane-changing model mobil for car-following models," *Transportation Research Record: Journal of the Transportation Research Board*, , no. 1999, pp. 86–94, 2007.

[19] M. Graf, O. Speidel, and K. Dietmayer, "A model based motion planning framework for automated vehicles in structured environments," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, June 2019, pp. 201–206.

[20] O. Speidel, M. Graf, A. Kaushik, T. Phan-Huu, A. Wedel, and K. Dietmayer, "Trajectory planning for automated driving in intersection scenarios using driver models," in *2020 5th International Conference on Robotics and Automation Engineering (ICRAE)*, 2020, pp. 131–138.

[21] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*, Malaysia; Pearson Education Limited,, 2016.

[22] S. Bhattacharya, "Discrete optimal search library (dosl): A template-based c++ library for discrete optimal search," 2017, Available at https://github.com/subh83/DOSL.

PERCEPTION THROUGH 2D-MIMO FMCW AUTOMOTIVE RADAR UNDER ADVERSE WEATHER

Xiangyu Gao, Sumit Roy, Guanbin Xing, Sian Jin

{xygao, sroy, gxing, sianjin} @uw.edu
Department of Electrical and Computer Engineering, University of Washington

ABSTRACT

Millimeter-wave (mmWave) radars are being increasingly integrated in commercial vehicles to support new Adaptive Driver Assisted Systems (ADAS) features that require accurate location and Doppler velocity estimates of objects, independent of environmental conditions. To explore radar-based ADAS applications, we have updated our test-bed with Texas Instrument's mmWave cascaded FMCW radar (TIDEP-01012) that forms a non-uniform 2D MIMO virtual array. In this paper, we develop the necessary received signal models for applying different direction of arrival (DoA) estimation algorithms and experimentally validating their performance on formed virtual array under controlled scenarios. To test the robustness of mmWave radars under adverse weather conditions, we collected raw radar dataset (I-Q samples post demodulated) for various objects by a driven vehicle-mounted platform, specifically for snowy and foggy situations where cameras are largely ineffective. Initial results from radar imaging algorithms to this dataset are presented.

Index Terms— mmWave, FMCW, 2D MIMO, DoA, non-uniform array, robustness, adverse weather.

1. INTRODUCTION

To meet requirements for ADAS and especially L4/L5 autonomous driving [1], automotive radars need to have a high angular resolution. There are several ways of improving radar angular resolution: 1) using more physical antenna elements, or a non-uniform array with larger antenna distance; 2) increasing the antenna aperture via synthetic aperture radar [2] concepts that exploit vehicle-mounted radar movement; 3) forming virtual array via multiple-input and multiple-out (MIMO) radar operations [3, 4]. In MIMO radar, multiple transmit (TX) antennas send orthogonal signals, which enables the contribution of each TX signal to be extracted at each receive (RX) antenna. Hence a physical TX array with M_{T} elements and RX array with M_{R} elements will result in a virtual array with upto $M_{\mathrm{T}}M_{\mathrm{R}}$ unique (non-overlapped) virtual elements [5]. To reduce array cost (fewer physical antenna elements), non-uniform arrays spanning large apertures, e.g., minimum redundancy array (MRA) [6] have been proposed.

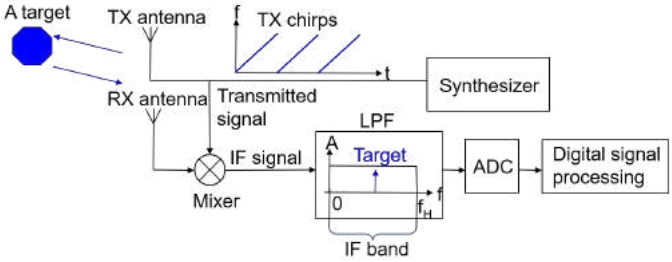


Fig. 1. Basic FMCW radar block diagram

From the received signals at RX array elements, the DoA of targets can be extracted by proper signal processing. The FFT-based DoA method and the multiple signal classification (MUSIC) are discussed and experimented in [7] on a 2 TX and 4 RX MIMO radar. Compressive sensing (CS) methods have been exploited in DoA estimation to exploit the inherent spatial sparsity of targets, via recovery algorithms from spatially under-sampled measurements [8]. For our application, CS-based DoA is applied to *non-uniformly spaced array*, that potentially enable higher resolution by reconstructing the observations at the missing elements [9]. Further, CS is known to mitigate the high sidelobes originating from non-uniform array [10] and therefore reduce false alarms [11, 12].

While mmWave radars are generally known for excellent environmental robustness under adverse weather conditions [13], there has been little published studies to date experimentally verifying this hypothesis due to the difficulty of capturing such data. Prior work [14] studied the effect of fog on the mmWave propagation, and Gao et al. [15] showed a robust and high-performance object recognition algorithm verified on nighttime data where cameras are largely ineffective. In this paper, we present a new CS-based DoA algorithm, whose performance is validated for data obtained using a frequency-modulated continuous wave (FMCW) 77 GHz radar test platform that enables a non-uniform 2D MIMO virtual array. In addition, we present some initial results regarding operational robustness to inclement weather by stress-testing performance of DoA on a collected dataset for snowy and foggy conditions.

2. FMCW MIMO RADAR

2.1. FMCW Radar and Range Estimation

FMCW radar transmits periodic wideband linear frequency-modulated (LFM, also called chirps) signal as shown in Fig. 1.

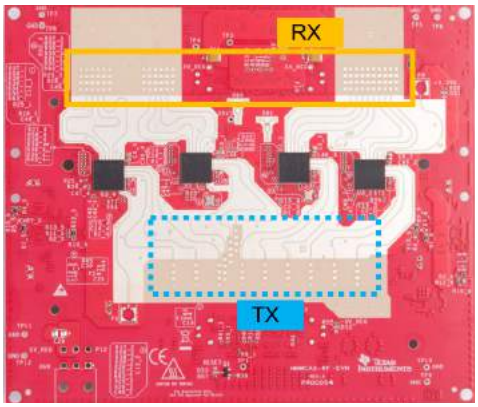


Fig. 2. Texas Instrument 4-chip cascaded radar board [16] and the position of antennas.

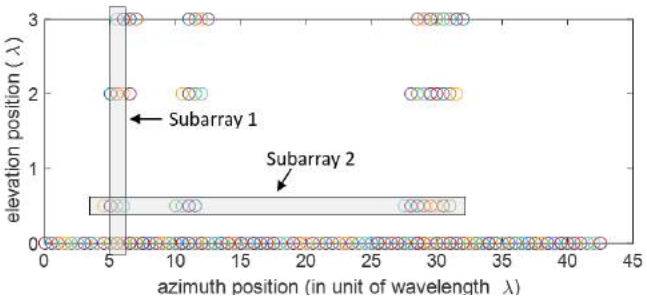


Fig. 3. 2D MIMO virtual array formed by 12 TX and 16 RX.

The TX signal is reflected from targets and received at the radar receiver. FMCW radars can detect targets' range and velocity from the RX signal using the stretch or de-chirping processing structure [7] in Fig. 1. A mixer at the receiver multiplies the RX signal with the TX signal to produce an intermediate frequency (IF) signal. Since the RX and the TX signal are both LFM signal with constant frequency difference determined by target's location, the IF signal is a single-tone signal. For example, the IF signal for a target at range r has frequency $f_{\text{IF}} = \frac{2r}{c}S$, the multiplication of round-trip delay $\frac{2r}{c}$ with chirp slope S, where c denotes the speed of the light. Thus, detecting the frequency of the IF signal can solve the target range. At the end of receiver, IF signal is passed into an anti-aliasing low-pass filter (LPF) and an analog-to-digital converter (ADC) for following digital signal processing. A fast Fourier transform (FFT) is widely adopted to estimate f_{IF} to infer r, and hence such operation is called the *Range FFT*.

2.2. MIMO Radar and Virtual Array

To estimate the direction of angle of targets relative to a receiver orientation, an antenna array is needed. In MIMO radar, a virtual array located at the spatial convolution of TX antennas and RX antennas is enabled by the orthogonality of TX signal [17]. The convolution produces a set of virtual element locations that is the sum of the TX and RX element locations. For example, if an automotive radar consists of a RX linear array of M_{R} elements with $\lambda/2$ spacing combined with a TX array of 2 elements which are spaced $M_{\text{R}}\lambda/2$ apart, the synthesized MIMO virtual array is a $2M_{\text{R}}$-elements uniform linear array (ULA) with $\lambda/2$ spacing.

We adopt TIDEP-01012 [16], a high-resolution mmWave FMCW radar board composed of four AWR2243 chips from Texas Instrument (TI) for experiments. This radar includes 12 TX and 16 RX antennas placed in specific 2D manner shown in Fig. 2, which creates a 2D virtual array (Fig. 3) with 192 elements via the spatial convolution of all TX and RX. The resulting virtual array has some overlapped elements and is mostly sparse except the bottom row (a ULA with 86 elements). For processing, we selected data for 2 subarrays - the vertical subarray 1 is a MRA [6] with 4 non-uniform spacing elements spanning 3λ aperture, and the non-uniform horizontal subarray 2 with 16 elements spanning 26.5λ.

3. SYSTEM MODEL FOR DOA ESTIMATION

Without loss of generality, we consider a RX uniform plane array (UPA) in the vertical plane with $M(N)$ antenna elements in each row (column), respectively. The array response is given by [18]:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{n}$ is a noise term, $\mathbf{x} = [\beta_1, \ldots, \beta_K]^{\text{T}}$ is the reflection coefficient matrix for K targets, and $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_K]$ is the array steering matrix with

$$\mathbf{a}_k = \mathbf{a}(u_k) \otimes \mathbf{a}(v_k)$$

$$\mathbf{a}(u_k) = \left[1, e^{j\frac{2\pi}{\lambda}d\sin\phi_k}, \ldots, e^{j\frac{2\pi}{\lambda}(N-1)d\sin\phi_k}\right]^{\text{T}}$$

$$\mathbf{a}(v_k) = \left[1, e^{j\frac{2\pi}{\lambda}d\sin\theta_k\cos\phi_k}, \ldots, e^{j\frac{2\pi}{\lambda}(M-1)d\sin\theta_k\cos\phi_k}\right]^{\text{T}}$$

Here, $\mathbf{a}(u_k)$ and $\mathbf{a}(v_k)$ are steering vectors for elevation angle ϕ_k and azimuth angle θ_k for kth target, respectively. $\otimes$ denotes the Kronecker product operation, and d is the antenna spacing. For a 1D ULA, angle finding can be done with digital beamforming by performing FFT across the received signal of array elements [7]. This FFT-based method can be extended to above 2D UPA, i.e., perform first FFT on the horizontal elements and second FFT on the elevation elements, which is *computationally efficient* but has low resolution.

3.1. MUSIC

MUSIC belongs to the class of *eigen-decomposition* based DoA estimators that construct the $(MN-K)$-dimension noise subspace U_{n} and the left K-dimension signal subspace from the covariance matrix of received signals $\mathbf{y}$ [19]. The azimuth

[1] θ_{res} is determined by maximum horizontal aperture length $L_{\text{h}} = 42.5\lambda$.
[2] ϕ_{res} is determined by the maximum vertical aperture length $L_{\text{v}} = 3\lambda$.
[3] T_{c} is equal to chirp interval times number of TX antennas.

Table 1. Parameter calculation (based on [7]) and configuration for 4-chip cascaded radar test-bed

Parameter	Calculation Equation	Configuration	Value
Range resolution (R_{res})	$R_{\text{res}} = \frac{c}{2B} = 0.39\,\text{m}$	Frequency (f_{c})	77 GHz
Velocity resolution (V_{res})	$V_{\text{res}} = \frac{\lambda}{2N_{\text{c}}T_{\text{c}}} = 0.0631\,\text{m/s}$	Sweep Bandwidth (B)	384 MHz
Azimuth Angle resolution (θ_{res}) [1]	$\theta_{\text{res}} = \frac{\lambda}{L_{\text{h}}\cos\theta} \approx 1.35°$	Sweep slope (S)	45 MHz/µs
Elevation Angle resolution (ϕ_{res}) [2]	$\phi_{\text{res}} = \frac{\lambda}{L_{\text{v}}\cos\theta} \approx 19°$	Sampling frequency (f_{s})	15 Msps
Max operating range (R_{max})	$R_{\text{max}} = \frac{f_{\text{s}}c}{2S} = 50\,\text{m}$	Num of chirps in one frame (N_{c})	128
Max operating velocity (V_{max})	$V_{\text{max}} = \frac{\lambda}{4T_{\text{c}}} = 4.04\,\text{m/s}$	Num of samples of one chirp (N_{s})	128
		Duration of chirp [3] and frame (T_{c}, T_{f})	240 µs, 1/30 s

and elevation angles (θ_k, ϕ_k) of the kth target can be found as peak on 2D MUSIC spectrum, which is given by [19]:

$$P_{\text{MUSIC}}(\theta_k, \phi_k) = \frac{1}{\mathbf{a}_k^{\text{H}} \boldsymbol{U}_{\text{n}} \boldsymbol{U}_{\text{n}}^{\text{H}} \mathbf{a}_k} \tag{2}$$

3.2. Compressive Sensing (CS)

To apply CS to DoA estimation, we need to define a search grid of K_{g} $\left(K_{\text{g}} \gg K\right)$ potential incident angles, and construct an hypothetical array steering matrix $\tilde{\mathbf{A}} = \left[\mathbf{a}_1, \ldots, \mathbf{a}_{K_{\text{g}}}\right]$ and the reflection coefficient matrix $\tilde{\mathbf{x}} = \left[\beta_1, \ldots, \beta_{K_{\text{g}}}\right]^{\text{T}}$.

The CS-based DoA estimation problem can be solved by an ℓ_1-norm regularized convex optimization, named square-root LASSO [8]:

$$\min_{\tilde{\mathbf{x}}} \xi \|\tilde{\mathbf{x}}\|_1 + \|\tilde{\mathbf{A}}\tilde{\mathbf{x}} - \mathbf{y}\|_2 \tag{3}$$

where $\|\cdot\|_1$ is the ℓ_1-norm forces the sparsity constraint, and $\xi > 0$ is a regularization parameter.

Above MUSIC and CS estimator are modeled for 2D UPA, and thus can address azimuth and elevation DoA estimation together. For ease of performing experiments and testing performance, we only use them for 1D DoA estimation next.

4. EXPERIMENTS

4.1. Radar Test-bed and Configuration

We assembled a test-bed (see Fig. 4) with the TIDEP-01012 radar [16] and binocular FLIR cameras (left and right). Binocular cameras are synchronized with radar to provide the visualization for the imaging scenarios. The 4-chip cascaded radar forms a large 2D-MIMO virtual array (see Fig. 3) via the time-division multiplexing (TDM) [3] on 12 TX antennas, resulting in substantial raw data ($\sim 378\,\text{MB}$) per second. Other configuration values of this radar are shown in Table. 1.

4.2. DoA Estimation on Non-uniform Array

Since the virtual array in Fig. 3 is mostly non-uniform, we evaluate the performance of different DoA estimators with

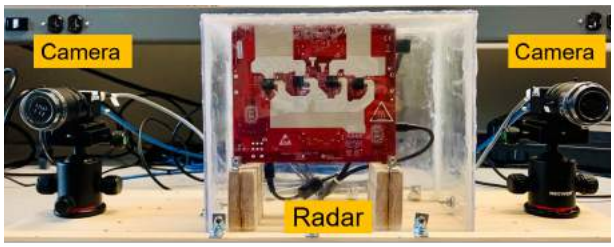


Fig. 4. 4-chip cascaded radar test-bed with 2 cameras.

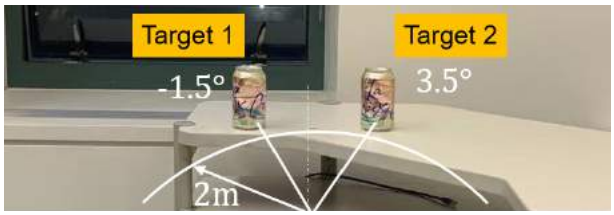


Fig. 5. Experiment setup with two targets separated by 5°.

it. First, we simulate the radar received signal of vertical subarray 1 (a MRA in Fig. 3) for two point targets located at range 20 m, $-6°$ and 5° elevation respectively. Three DoA algorithms - FFT, MUSIC, and CS - are implemented on the simulated signal to obtain the spectrum maps shown in Fig. 6. Note that we choose regularization parameter $\xi = 1.4$ in CS reconstruction, based on exhaustive search. The results show that MUSIC and CS-based DoA estimators achieve *improved resolution* for non-uniform array by the ability to separate two targets, while FFT does not. Besides, CS generates the sparse solution that *avoids high sidelobes* at around $\pm 30°$. To compare the DoA estimation accuracy, we calculate the root-mean-square errors (RMSE) for all methods by averaging over 30 simulation rounds. We got the RMSE of MUSIC method (2.4609°) and CS method (0.3162°), which demonstrates that CS-based DoA estimation is more accurate than MUSIC on selected non-uniform linear array.

Second, we employed a setup with two close targets placed at 2 m, $-1.5°$ and 3.5° azimuth respectively (see Fig. 5), and collected the real radar return signal. The power spectrum

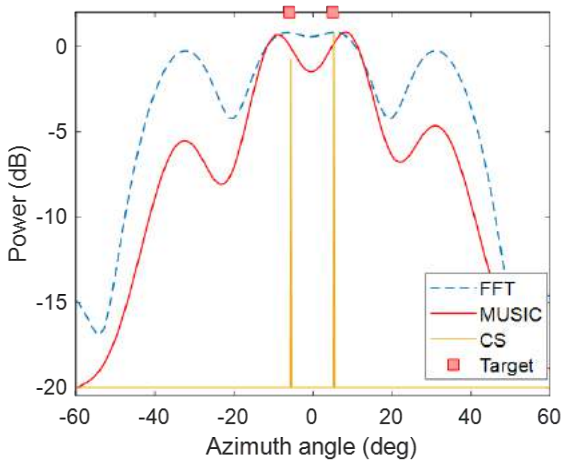


Fig. 6. DoA spectrums for FFT, MUSIC, and CS estimators with simulation on vertical non-uniform subarray 1 (in Fig. 3).

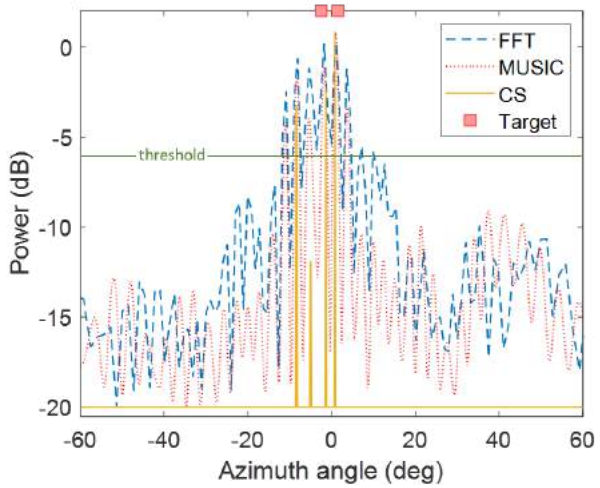


Fig. 7. DoA spectrums for FFT, MUSIC, and CS estimators with experiment on horizontal subarray 2 (in Fig. 3).

for the received signal for horizontal subarray 2 (in Fig. 3) using our CS-DOA approach is shown in Fig. 7 and shows 3 strong peaks - two of them corresponds to targets, while the third is likely a spurious reflection from an indoor wall. To evaluate the false alarms of FFT and MUSIC based DoA estimation caused by non-uniform array spacing, we set a $-6\,\mathrm{dB}$ threshold and count the additional peaks exceeding magnitude threshold. Results show that CS, FFT and MUSIC estimator have 0, 5 and 3 false alarms respectively, which verifies that CS is more robust for DoA estimation. It is to be noted that the performance of CS is dependant on the choice of regularization value ξ. The optimal regularization ξ is a function of the number of targets, and may be determined by exhaustive search to find the optimal value.

Fig. 8. Test-bed mounted on a vehicle for dataset collection.

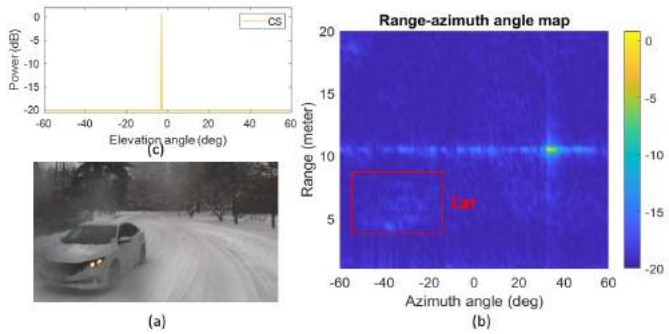


Fig. 9. An example of adverse-weather radar dataset: a moving vehicle with its (a) camera image, (b) range-azimuth angle map, (c) elevation DoA spectrum for range $r = 4.3\,\mathrm{m}$.

4.3. Radar Dataset and Imaging for Adverse Weathers

For promoting the development of high-level radar ADAS applications (e.g., object recognition [15]) under critical adverse weathers, we mount the test-bed on a vehicle (see Fig. 8) and collect raw radar I-Q samples dataset for various objects (pedestrian, cyclist, and car) by driving vehicle in *snowy and foggy conditions* where camera images are compromised.

We present an example from the collected dataset in Fig. 9, with a camera image of a moving car and corresponding radar imaging results. The range-azimuth angle map (see Fig. 9(b)) is generated by performing Range FFT and FFT-based DoA on the radar data of bottom-row ULA (in Fig. 3). We also show the elevation DoA spectrum for range $4.3\,\mathrm{m}$ (in Fig. 9(c) with ground truth around $-3°$), which is obtained by executing CS on corresponding radar data of vertical subarray 1 in Fig. 3. According to qualitative results, the vehicle object is still visible in radar image even with the attenuation from snow and fog, which validates the robustness of mmWave radar primitively.

5. CONCLUSION

A high-resolution mmWave FMCW radar test-bed with non-uniformly spaced 2D-MIMO virtual array was used for testing a CS-DoA algorithm performance, initially calibrated and benchmarked for some test cases followed by evaluation for a dataset collected under adverse weather (snow, fog) conditions.

6. REFERENCES

[1] NHTSA, "Automated vehicles for safety," https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety.

[2] X. Gao, S. Roy, and G. Xing, "Mimo-sar: A hierarchical high-resolution imaging algorithm for fmcw automotive radar," 2021, Available at https://arxiv.org/pdf/2101.09293.pdf.

[3] Sandeep Rao, *White paper: MIMO Radar*, Number SWRA554A. Texas Instrument, 2017, Available at https://www.ti.com/lit/an/swra554a/swra554a.pdf.

[4] Guohua Wang, Siddhartha, and Kumar Vijay Mishra, "Stap in automotive mimo radar with transmitter scheduling," in *2020 IEEE Radar Conference (RadarConf20)*, 2020, pp. 1–6.

[5] W. Wang, "Virtual antenna array analysis for mimo synthetic aperture radars," *International Journal of Antennas and Propagation*, vol. 2012, pp. 1–10, 2012.

[6] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Transactions on Antennas and Propagation*, vol. 16, no. 2, pp. 172–175, 1968.

[7] X. Gao, G. Xing, S. Roy, and H. Liu, "Experiments with mmwave automotive radar test-bed," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 1–6.

[8] Zai Yang, Jian Li, Petre Stoica, and Lihua Xie, "Chapter 11 - sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing, Volume 7*, Rama Chellappa and Sergios Theodoridis, Eds., pp. 509–581. Academic Press, 2018.

[9] F. Roos, P. Hügler, L. L. T. Torres, C. Knill, J. Schlichenmaier, C. Vasanelli, N. Appenrodt, J. Dickmann, and C. Waldschmidt, "Compressed sensing based single snapshot doa estimation for sparse mimo radar arrays," in *2019 12th German Microwave Conference (GeMiC)*, 2019, pp. 75–78.

[10] M. Andreasen, "Linear arrays with variable interelement spacings," *IRE Transactions on Antennas and Propagation*, vol. 10, no. 2, pp. 137–143, 1962.

[11] A. Correas-Serrano and M. A. González-Huici, "Experimental evaluation of compressive sensing for doa estimation in automotive radar," in *2018 19th International Radar Symposium (IRS)*, 2018, pp. 1–10.

[12] F. Roos, P. Hügler, J. Bechter, M. A. Razzaq, C. Knill, N. Appenrodt, J. Dickmann, and C. Waldschmidt, "Effort considerations of compressed sensing for automotive radar," in *2019 IEEE Radio and Wireless Symposium (RWS)*, 2019, pp. 1–3.

[13] Shizhe Zang, Ming Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and Mohamed Ali Kâafar, "The impact of adverse weather conditions on autonomous vehicles: How rain, snow, fog, and hail affect the performance of a self-driving car," *IEEE Vehicular Technology Magazine*, vol. 14, pp. 103–111, 2019.

[14] Yosef Golovachev, Ariel Etinger, Gad Pinhasi, and Yosef Pinhasi, "Millimeter wave high resolution radar accuracy in fog conditions-theory and experimental verification," *Sensors*, vol. 18, 07 2018.

[15] X. Gao, G. Xing, S. Roy, and H. Liu, "Ramp-cnn: A novel neural network for enhanced automotive radar object recognition," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 5119–5132, 2021.

[16] Texas Instrument, *White paper: Imaging Radar Using Cascaded mmWave Sensor Reference Design*, Number TIDUEN5A. Texas Instrument, 2019, Available at https://www.ti.com/lit/ug/tiduen5a/tiduen5a.pdf.

[17] S. Sun, A. P. Petropulu, and H. V. Poor, "Mimo radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 98–117, 2020.

[18] K. Goto, T. Akao, K. Maruta, and C. Ahn, "Reduced complexity direction-of-arrival estimation for 2d planar massive arrays: A separation approach," in *2018 18th International Symposium on Communications and Information Technologies (ISCIT)*, 2018, pp. 48–53.

[19] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

DIFFERENCE CO-CHIRPS-BASED NON-UNIFORM PRF AUTOMOTIVE FMCW RADAR

Lifan Xu[1], *Shunqiao Sun*[1] *and Kumar Vijay Mishra*[2]

[1]Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL 35487
[2]United States CCDC Army Research Laboratory, Adelphi, MD 20783

ABSTRACT

We propose an automotive radar system that transmits at non-uniform pulse repetition frequency (PRF) to achieve high-resolution range and Doppler estimation while transmitting sparsely along slow-time following the difference co-chirps schemes, e.g., coprime and nested chirps. At the receiver, the radar admits undersampled slow-time signals for Doppler estimation. In a single coherent processing interval (CPI), the missing Doppler samples along slow-time are interpolated via a Doppler covariance matrix that is constructed using fast-time samples. Our *co*-chirp joint range-*D*oppler estimation with *D*opp*ler* de-aliasing (CoDDler) algorithm jointly estimates the range and Doppler. The Doppler spectrum obtained from the interpolated Doppler samples are utilized to de-aliase any false Doppler peaks in the sparse estimation. The proposed non-uniform PRF automotive radar provides the possibility for transmission coordination in a time division multiplexing fashion to avoid mutual interference by saving nearly 88% of time-on-target.

***Index Terms*—** Automotive radar, difference co-chirps, FMCW, non-uniform PRF, sparse reconstruction.

1. INTRODUCTION

Automotive radar systems exploit frequency modulated continuous wave (FMCW) signals at millimeter-wave frequencies to enable high-resolution target range and velocity estimation while using a lower cost and simpler hardware than a light detection and ranging (LiDAR) system [1–4]. As more vehicles are equipped with automotive radar operating within the same 76-81 GHz frequency band, the mutual interference among automotive radars becomes severe. The mutual interference will decrease the detection performance of automotive radar. Many techniques have been proposed to mitigate the interference. The simplest way is to notch out the corrupted samples caused by the interference [3] or transmit multi-purpose waveforms [5, 6]. An alternative is to schedule the radar operation time such that mutual coherence is avoided [7, 8]. However, this strategy would result in loss of useful dwell time on targets. This is addressed through transmitting chirps sparsely along slow-time and then employ sparse reconstruction techniques to recover the targets [9–11]. Such a strategy has potential to coordinate the transmission among multiple automotive radars to avoid or greatly reduce mutual interference without loss of performance.

In this context, radar waveforms that transmit at non-uniform pulse repetition frequency (PRF) have been developed in [10, 12–14]. In [12], a weight interpolation technique was considered to handle the high sidelobes in the Doppler spectrum caused by non-uniform pulsing. In [13], after interpolation to suppress these sidelobes, the non-uniform pulses are processed via non-uniform fast Fourier transform (NUFFT). A direct interpolation of Fourier coefficients is avoided in [10] by employing the compressed sensing (CS) technique to recover the Doppler information. Optimal pulse transmission structure and sampling rules were considered in [14] to control the sidelobe level of Doppler spectrum.

In this context, there is a rich heritage of research on non-uniform sampling in the spatial domain, e.g., using sparse arrays [15, 16]. For example, difference coarray concept has been exploited for direction of arrival (DOA) estimation in passive sensing. Some well-known difference coarrays include the minimum redundancy array (MRA) [17], nested array [18], coprime array [19–21], and super nested array [22, 23]. When a sufficient number of snapshots are available, the difference coarray concept is utilized to construct a coarray with significantly increased virtual sensors from a sparse physical array [24]. Difference coarray concept has been utilized for spectrum estimation of a wide sense stationary (WSS) process with significantly reduced sampling rate [25]. In the automotive sensing scenario, however, the environment is highly dynamic [26]. The positions of radar-mounted vehicle and objects often change rapidly [3].

In this paper, we exploit the concept of difference coarray for automotive radar waveform design. We propose *difference co-chirps* for an FMCW radar, where a coherent processing interval (CPI) comprises a sparse chirp sequence following the co-chirp concept instead of a a consecutive chirp sequence along slow-time with a high PRF. At the receiver, we sparsely sample the radar signals along slow-time. In order to jointly recover the range and Doppler, we adopt a two-dimensional CS (2D-CS) algorithm followed by a Doppler de-aliasing step. The missing samples along the slow-time are filled via the sampling covariance matrix that is constructed using the fast-time samples. The filled slow-time samples are utilized to de-aliasing any false Doppler peaks in the CS estimation. Our proposal does not rely on multiple snapshots from multiple CPIs. Furthermore, it allows the transmission coordination among multiple automotive radars to avoid potential mutual interference.

The rest of the paper is organized as follows. In the next section, we introduce the conventional uniform PRF and our proposed co-chirp FMCW radars. In Section 3, we describe our *co*-chirp joint range-*D*oppler estimation with *D*opp*ler* de-aliasing (CoDDler) procedure for the co-chirp radar. We validate our models and methods through numerical experiments in Section 4. We conclude in Section 5. Throughout this paper, upper-case and lower-case bold characters denote matrices and vectors respectively. Matrix vectorization operation is denoted by $\mathrm{vec}(\cdot)$. The conjugate transpose is $(\cdot)^H$. The complex values set is $\mathbb{C}$. The ceiling operation is denoted by $\lceil\cdot\rceil$.

2. SYSTEM MODEL

We briefly describe the conventional FMCW radar and follow with the difference co-chirp system.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

2.1. State-of-the-art FMCW radar

Consider a monostatic FMCW radar that transmits a linear frequency ramp. The transmit signal for one ramp at m-th chirp with bandwidth B and the duration time T is

$$s(m,t) = \text{rect}\left(\frac{t-mT_p}{T}\right) e^{\text{j}2\pi\left[f_c+\frac{B}{T}(t-mT_p)\right](t-mT_p)}, \quad (1)$$

where T_p denotes the pulse repetition interval (PRI) and

$$\text{rect}\left(\frac{t-\tau}{T}\right) = \begin{cases} 1, & \tau \le t \le \tau+T \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The phase of the transmit signal $s(m,t)$ is

$$\begin{aligned}\varphi_T\left(t-mT_p\right) =& 2\pi \int_{mT_p}^{mT_p+t} \left[f_c + \frac{B}{T}(t-mT_p)\right] dt \\ =& 2\pi\left(f_c t + \frac{1}{2}\cdot\frac{B}{T}t^2\right) - \varphi_{T_0}, \end{aligned} \quad (3)$$

where φ_{T_0} is the initial phase. We consider K_u noncoherent signals and P_c coherent signals in far field. The delayed version of transmit signal is the noiseless received signal

$$\begin{aligned} y(t) = &\sum_{k_u=1}^{K_u} \alpha_{k_u} e^{\text{j}2\pi\left[f_c(t-\tau_{k_u})+\frac{B}{2T}(t-\tau_{k_u})^2\right]} \\ &+ \sum_{p_c=1}^{P_c} \alpha_{p_c} e^{\text{j}2\pi\left[f_c(t-\tau_{p_c})+\frac{B}{2T}(t-\tau_{p_c})^2\right]}, \end{aligned} \quad (4)$$

where α_{k_u} (τ_{k_u}) and α_{p_c} (τ_{p_c}) denotes the reflection coefficients (time delays) of noncoherent and coherent signals, respectively.

The above received signal is first de-chirped with the transmit signal. The de-chirped signal is called beat signal, whose phase is

$$\Delta\varphi(t) = \varphi_T(t) - \varphi_T(t-\tau_k) = 2\pi\left(f_c\tau_k + \frac{B}{T}t\tau_k - \frac{B}{2T}\tau_k^2\right), \quad (5)$$

where τ_k denotes the delay between the transmitted and received signal of k-th target. The square term of time τ_k in the equation (5) is negligible in short range automotive radars since it typically holds that $\tau_k/T \ll 1$. Here, in the de-chirped signal, t starts from zero for each chirp.

Assume the k-th target has range of r_k with constant velocity v_k. Then, the round-trip transmission delay for the k-th target is $\tau_k = 2(r_k + v_k t)/c$, where c is the speed of light. The phase of the de-chirp signal is

$$\Delta\varphi(t) = 2\pi\left[\frac{2f_c r_k}{c} + \left(\frac{2f_c v_k}{c} + \frac{2Br_k}{cT}\right)t + \frac{2Bv_k}{cT}t^2\right], \quad (6)$$

where $\frac{2Bv_k}{Tc}t^2$ is negligible in typical automotive radars and $\frac{2f_c r_k}{c}$ is a constant phase that can be absorbed into the reflection coefficients α_{k_u} or α_{p_c}. The beat frequency of the k-th target is $f_b^k = f_R^k + f_D^k$, where $f_R^k = \frac{2Br_k}{Tc}$ and $f_D^k = \frac{2f_c v_k}{c}$ are, respectively, range and Doppler frequencies of the k-th target.

In automotive radars with maximum detection range of hundreds of meters, it holds that $f_b \ll B$. As a result, the beat signals can be sampled using a much cheaper low-rate analog-to-digital converter (ADC). Assume T_A denotes the sampling interval in fast-time and $1/T_A > 2f_b^{\max}$, where $f_b^{\max}$ denotes the maximum beat frequency. The i-th sample in the m-th chirp is

$$\begin{aligned} y(m,i) = &\sum_{k_u=1}^{K_u} \alpha_{k_u} e^{\text{j}2\pi\left(f_b^{k_u} iT_A + f_D^{k_u} mT_p\right)} + \\ &\sum_{p_c=1}^{P_c} \alpha_{p_c} e^{\text{j}2\pi\left(f_b^{p_c} iT_A + f_D^{p_c} mT_p\right)}. \end{aligned} \quad (7)$$

Assume a CPI consists of M chirps and number of samples in each chirp is I. The sampled automotive radar data cube is denoted by $\mathbf{Y} \in \mathbb{C}^{I\times M}$ whose entries are $y(m,i)$.

For a typical automotive radar, it holds that $f_D \ll f_R$. Thus, the Doppler frequency f_D is negligible in a single chirp and range is estimated by applying fast Fourier transform (FFT) along fast-time samples in the above-mentioned data cube. For each range bin, the range frequency f_R is constant across the slow-time. Thus, the Doppler is estimated by applying FFT along the slow-time in data cube $\mathbf{Y}$ [3]. To avoid ambiguity in Doppler spectrum estimation in uniform PRF radar, it is required that $f_{\text{PRF}} \ge 2f_D^{\max}$, where $f_{\text{PRF}} = 1/T_p$ is the PRF of chirps and $f_D^{\max}$ denotes the maximum Doppler frequency.

In this paper, instead of transmitting a uniform chirp sequence with a high PRF, we propose an automotive radar system that transmits a non-uniform chirp sequence in each CPI following the co-chirp concept, such as coprime chirps and nested chirps. The challenge lies in that the non-uniform PRF violates the Nyquist sampling rate for Doppler estimation thereby leading to high sidelobes in the Doppler spectrum. Our goal is to develop high-resolution algorithm to jointly estimate range and Doppler under the difference co-chirps concept while avoiding Doppler ambiguity.

2.2. Difference co-chirp based FMCW radar

Consider a chirp set $\mathbb{S}_1 = \{m_1, m_2, ..., m_M\}$, which has M chirp entries. Let m_i denote the i-th chirp. The set of difference chirp indices is

$$\mathbb{S}_{\text{diff}} = \{m_i - m_j\}, \qquad \forall i,j \in \mathbb{S}_1 \quad (8)$$

In this difference co-chirps set, the entries occur only once.

2.2.1. FMCW radar with coprime chirps scheme

Consider an FMCW radar that transmits along the slow-time according to coprime chirps relationship. Two coprime numbers N_1 and N_2 are used to define a chirps slow-time slot set as

$$\begin{aligned} &\mathbb{S}_{\text{coprime}} = \\ &\{N_1 n_2, 0 \leqslant n_2 \leqslant N_2 - 1\} \cup \{N_2 n_1, 0 \leqslant n_1 \leqslant N_1 - 1\}. \end{aligned} \quad (9)$$

A FMCW chirp will be transmitted at the slow-time indices specified in the above set. The difference co-chirps set is

$$\mathbb{S}_{\text{diff}} = \{s_1 - s_2 | s_1, s_2 \in \mathbb{S}_{\text{coprime}}\}. \quad (10)$$

However, the difference co-chirps set does not include consecutive chirps from time slots $-N_2(N_1-1)$ to $N_1(N_2-1)$, and certain chirp indices are missing (see Fig. 1(b)).

2.2.2. FMCW radar with nested chirps scheme

We now examine the FMCW radar that sends pulses along the slow-time following the nested chirps relationship. Two-level chirp indices are used in nested chirps scheme. Specifically, the first and second levels consist of N_1 and N_2 chirps with corresponding PRIs as T_p and $(N_1+1)T_p$, respectively. Under the nested chirps scheme,

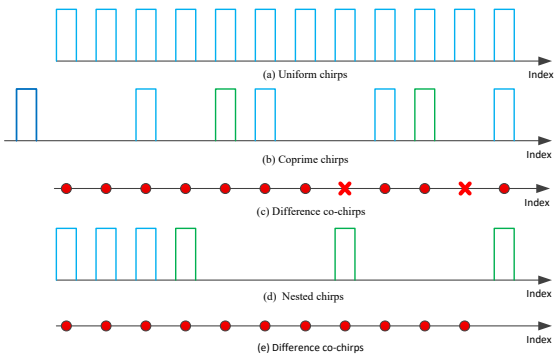


Fig. 1. Different non-uniform chirps schemes and their difference co-chirps.The missing chirp indices are denoted by $\times$, while the filled difference chirp indices are denoted by $\bullet$. (a) The uniform PRF chirps. (b) The coprime chirps scheme for $N_1 = 3, N_2 = 5$. (d) The nested chirps scheme for $N_1 = 3, N_2 = 3$.

FMCW radar transmit chirps at slow-time indices as per the set

$$\mathbb{S}_{\text{nested}} = \{1, 2, \cdots, N_1, (N_1+1), 2(N_1+1), \cdots, N_2(N_1+1)\}. \quad (11)$$

The set $\mathbb{S}_{\text{diff}} = \{n_1 - n_2 | n_1, n_2 \in \mathbb{S}_{\text{nested}}\}$ is called the difference of nested-chirps. The set containing unique chirp (UC) listed in $\mathbb{S}_{\text{diff}}$ is denoted as $\mathbb{S}_{\text{diff}}^{\text{UC}}$ (see Fig. 1(c)), where total transmitted chirps are $N = N_1 + N_2$. Under the nested scheme, the first N_1 transmitted chirps have PRI of T_p while the second N_2 chirps have PRI of $(N_1+1)T_p$. The sampled beat signal at the m-th chirps can be expressed in (7) for $m \in \mathbb{S}_{\text{nested}}$.

The Doppler velocity resolution is determined by the length of a CPI. To achieve the same Doppler resolution as uniform PRF scheme, FMCW radar under difference co-chirps schemes need to transmit over the whole CPI sparsely along the slow-time following the corresponding coprime or nested co-chirp rules. In Fig. 1, total 12 chirps need to be transmitted under the uniform PRF scheme in one CPI. On the other hand, for the same observation window time, only 8 and 6 chirps need to be transmitted under the coprime and nested chirp strategies, respectively.

3. JOINT RANGE AND DOPPLER ESTIMATION

We develop super-resolution algorithm CoDDler to jointly estimate the range and Doppler when FMCW chirps are transmitted sparsely along slow-time following the difference nested-chirps rule. Note that high sidelobes of the Doppler spectrum arising from the sparse sampling pose a challenge here. Our 2D-CS algorithm jointly estimates the range and Doppler using the sparse samples along slow-time. To remove the Doppler ambiguity, we devise a difference co-chirps interpolation based Doppler de-aliasing strategy.

Assume R_u and $v_{\max}$ denote the maximum detection range and velocity, respectively. To construct an appropriate CS dictionary [27, 28], the range and velocity are discretized into a fine grid with $M_r \times M_v$ points. The corresponding range and velocity grid sizes are $\frac{R_u}{M_r}$ and $\frac{2v_{\max}}{M_v}$, respectively. The ξ-th range and η-th discretized velocity are denoted as R_ξ and v_η. The corresponding beat frequency is $f_b^{\xi\eta} = f_R^{\xi} + f_D^{\eta}$. The corresponding constructed noise-free data matrix is denoted by $\mathbf{Z}_{\xi\eta} \in \mathbb{C}^{I \times N}$ whose elements are

$$z(n,i) = e^{\text{j}2\pi\left(f_b^{\xi\eta} i T_A + f_D^{\eta} n T_p\right)}, n \in \mathbb{S}_{\text{nested}}. \quad (12)$$

The dictionary of the 2D-CS is

$$\mathbf{A} = [\text{vec}(\mathbf{Z}_{11}), \cdots, \text{vec}(\mathbf{Z}_{1M_v}), \text{vec}(\mathbf{Z}_{21}), \cdots, \text{vec}(\mathbf{Z}_{M_r M_v})]. \quad (13)$$

In practice, the measurement is corrupted with noise, i.e., $\text{vec}(\mathbf{Y}) = \mathbf{A}\mathbf{x} + \mathbf{n}$, where $\mathbf{n}$ is the noise vector. Here, $\mathbf{x} \in \mathbb{C}^{M_v M_r \times 1}$ is a sparse vector, where $x_j = \alpha_h$ with $h = K_u$ or $h = P_c$ if the h-th target has range of $\frac{R_u}{M_r}\left\lceil \frac{j}{M_v} \right\rceil$ and velocity of $-v_{\max} + \frac{2v_{\max}}{M_v} \mod (j, M_r)$; otherwise, $x_j = 0$.

To range and Doppler, we solve the optimization problem

$$\text{minimize } \|\mathbf{x}\|_1 \quad \text{subject to } \|\text{vec}(\mathbf{Y}) - \mathbf{A}\mathbf{x}\|_2 \le \delta, \quad (14)$$

where δ is the noise bound. Some popular solvers such as Dantzig selector [29] or orthogonal matching pursuit (OMP) [30] are used to solve (14) to obtain the signal vector $\mathbf{x}$. Successful recovery of the sparse vector $\mathbf{x}$ requires that the dictionary matrix $\mathbf{A}$ has low value of its mutual coherence defined as

$$\mu(\mathbf{A}) = \max_{l \ne j} \frac{\mathbf{A}_l^{\text{H}} \mathbf{A}_j}{\|\mathbf{A}_l\|_2 \|\mathbf{A}_j\|_2}, \quad (15)$$

where $\mathbf{A}_j$ denotes the j-th column of matrix $\mathbf{A}$.

It follows that the sparse sampling along slow-time with difference co-chirps schemes leads to high sidelobes of Doppler spectrum resulting in high mutual coherence. Consequently, false Doppler detections are likely in the CS output. By utilizing the sparse samples from difference co-chirps, we interpolate the missing samples along slow-time for non-ambiguity Doppler estimation, which are then used for Doppler de-aliasing in the CS estimation output.

In each CPI, the missing samples along slow-time for Doppler estimation are interpolated via construction of a second-order covariance matrix, which requires a large number of snapshots. As mentioned earlier, the Doppler shift in a typical automotive radar is negligible during fast-time sampling of a single chirp and is viewed as a constant [3]. Therefore, we treat the fast-time samples as "snapshot" for Doppler covariance matrix construction. Following the equation (7), the i-th snapshot of slow-time samples or the i-th row of sparse radar data cube is

$$\mathbf{y}_{\text{nested}}^i = \mathbf{B}\mathbf{\Sigma}\mathbf{s}^i + \mathbf{n}^i, \quad (16)$$

where $\mathbf{B} = \left[\mathbf{b}(f_D^1), \mathbf{b}(f_D^2), \cdots, \mathbf{b}(f_D^K)\right] \in \mathbb{C}^{N \times K}$ is the Doppler manifold with $\mathbf{b}(f_D^k) = \left[e^{\text{j}2\pi f_D^k T_p}, \cdots, e^{\text{j}2\pi f_D^k N T_p}\right]^T$ and $\mathbf{s}^i = \left[e^{\text{j}2\pi f_b^k i T_A}, \cdots, e^{\text{j}2\pi f_b^k i T_A}\right]^T$. Here, $\mathbf{n}^i$ denotes the noise vector in the i-th snapshot of slow-time samples and $\mathbf{\Sigma} = \text{diag}([\alpha_1, \cdots, \alpha_K])$.

The missing Doppler samples along slow-time are interpolated via the Doppler autocorrelation $\mathbf{y}_{\text{nested}}^i \left(\mathbf{y}_{\text{nested}}^i\right)^H$, whose entries include $e^{\text{j}2\pi f_D^k (n_2 - n_1) T_p}$ for $n_1, n_2 \in \mathbb{S}_{\text{nested}}$, i.e., $e^{\text{j}2\pi f_D^k n T_p}$ for $n \in \mathbb{S}_{\text{diff}}$. From the nested-chirps properties, the indices in $\mathbb{S}_{\text{diff}}$ are consecutive. The sampling Doppler covariance matrix is

$$\mathbf{R}_{\text{nested}} = \frac{1}{I} \sum_{i=1}^{I} \mathbf{y}_{\text{nested}}^i \left(\mathbf{y}_{\text{nested}}^i\right)^H. \quad (17)$$

Let $\mathbf{d}_{\text{diff}}^{\text{UC}} = \text{unique}(\mathbf{R}_{\text{nested}})$ denote the averaged unique consecutive Doppler samples obtained from the sampling covariance matrix with indices defined in $\mathbb{S}_{\text{diff}}^{\text{UC}}$. Then, the Doppler spectrum is obtained by applying FFT on the interpolated Doppler samples along slow-time. The Doppler spectrum is accurate and robust, which indicates the targets' radar cross section (RCS). Thus, the Doppler spectrum is utilized to filter out false velocity peaks in the CS estimation. Algorithm 1 summarizes these steps.

Algorithm 1 *Co*-chirp joint range-*D*oppler estimation with *D*opp*ler* de-aliasing (CoDDler)

Input: N_1, N_2, M_v, M_r and the received sparse data cube $\mathbf{Y}$.
Output: de-aliasing CS range-Doppler spectrum.
Doppler spectrum with interpolated Doppler samples:

$\mathbf{R}_{\text{nested}} = \frac{1}{I}\sum_{i=1}^{I} \mathbf{y}^i_{\text{nested}} \left(\mathbf{y}^i_{\text{nested}}\right)^H$.

$\mathbf{d}^{\text{UC}}_{\text{diff}} = \text{unique}\left(\mathbf{R}_{\text{nested}}\right)$.

$\mathcal{D} = \text{FFT}\left\{\mathbf{d}^{\text{UC}}_{\text{diff}}\right\}$.

Range-Doppler estimation with 2D-CS and Doppler de-aliasing:

Discretize the range and velocity into a fine grid and construct dictionary matrix $\mathbf{A}$ according to (13).

Solve ℓ_1 norm optimization problem (14) by OMP.

Apply the Doppler spectrum $\mathcal{D}$ to filter out fake velocity peaks in CS estimation.

4. NUMERICAL RESULTS

We consider a two-level nested chirps FMCW radar with carrier frequency of f_c =77 GHz, and bandwidth of B = 150 MHz. The maximum unambiguous detection range is R_u = 200 m and the maximum unambiguous velocity is $v_{\max}$ = 63.89 m/s. The sampling frequency of beat signal is f_s = 27.3 MHz. The PRI is $T_p = 15.2\mu$s for uniform PRF automotive radar. In a CPI, the difference co-chirps based automotive radar first transmits N_1 = 17 chirps with PRI of T_p and then N_2 = 17 chirps with PRI of $(N_1+1)\,T_p$. Then, there are total 305 chirps in the interpolated Doppler samples. The signal-to-noise ratio (SNR) of the beat signal is set to 10 dB. Two targets at r_1 = 110 m, v_1 = 15 m/s and r_2 = 50 m, v_2 = 35 m/s with the same RCS are considered.

The range-Doppler spectrum obtained by applying the 2D-FFT on the sparse data with nested-chirps are shown in Fig.2(a). Note that there are many high sidelobes along Doppler axis because the sparse sampling along slow-time violates the Nyquist sampling criterion. The range-Doppler estimation on the sparse data with nested-chirps using 2D-CS is plotted in Fig. 2(b). The high coherence of dictionary $\mathbf{A}$ leads to false peaks in the velocity estimation, which cannot be eliminated without prior knowledge of targets.

Fig. 2(c) plots the Doppler spectrum obtained from the interpolated Doppler samples along slow-time, which clearly show two peaks at the ground-truth locations. The threshold obtained from the Doppler spectrum is used to filter out artifacts in the 2D-CS estimation. The 2D-CS estimation after de-aliasing using Doppler spectrum is shown in Fig. 2(d), where the false peaks are mitigated.

The range-Doppler spectrum obtained by applying the 2D-FFT on the sparse data with co-prime chirps are shown in Fig. 3(a), while the range-Doppler estimation with 2D-CS on the sparse samples are plotted in Fig. 3(b). Since the interpolated Doppler samples along slow-time are not consecutive under the difference co-prime chirps scheme, there are high sidelobes in the Doppler spectrum of the interpolated Doppler samples (Fig. 3(c)), which cannot be used for Doppler de-aliasing in the 2D-CS estimation.

In both examples, a uniform PRF radar would transmit 306 chirps in a CPI. On the contrary, the difference co-chirps schemes with nested chirps and coprime chirps schemes only need 34 and 37 chirps to achieve the same Doppler resolution, respectively, with 89% and 88% savings in the dwell time, respectively.

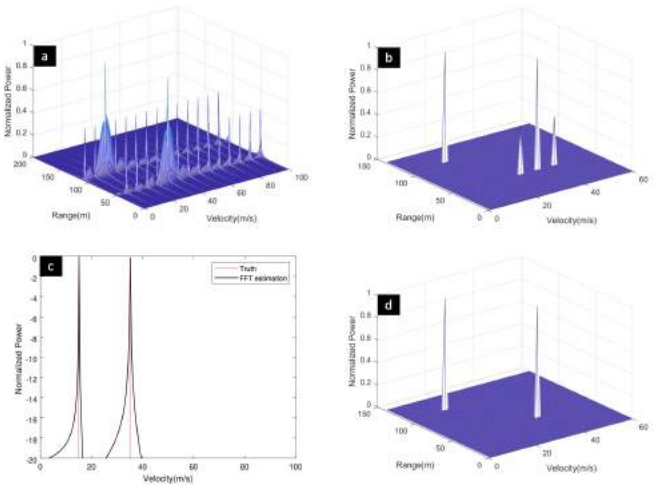
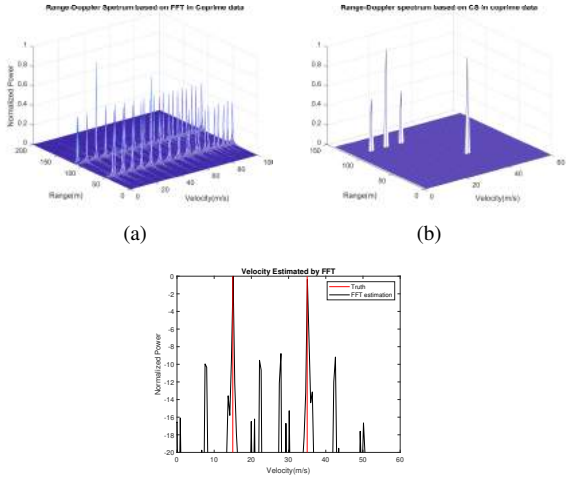


Fig. 2. Range-Doppler spectrum under the difference nested-chirps scheme. (a) 2D-FFT on the sparse samples. (b) 2D-CS on the sparse samples. (c) Doppler spectrum of the interpolated slow-time samples. (d) 2D-CS on the sparse samples after Doppler de-aliasing.



(a) (b)

(c)

Fig. 3. Range-Doppler spectrum under the difference co-prime chirps scheme. (a) 2D-FFT on the sparse samples. (b) 2D-CS on the sparse samples. (c) Doppler spectrum of the interpolated slow-time samples.

5. SUMMARY

We proposed difference co-chirps based non-uniform PRF automotive FMCW radar that exploits the difference coarray concepts. In this approach, the automotive radar transmits sparsely in a CPI with the equivalent co-chirp determining the chirp sequence. In particular, we investigated co-prime and nested chirp sequences. Our proposed CoDDleR jointly estimates the range and Doppler from the radar signals sparsely sampled along slow-time. Numerical results demonstrated the feasibility of the proposed method in achieving high-resolution range-Doppler estimation while mitigating false Doppler peaks. The saving in the dwell time with co-chirp radar is more than 88% when compared to the uniform PRF radar.

6. REFERENCES

[1] S. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 22–35, 2017.

[2] F. Engels, P. Heidenreich, A. M. Zoubir, F. Jondral, and M. Wintermantel, "Advances in automotive radar: A framework on computationally efficient high-resolution frequency estimation," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 36–46, 2017.

[3] S. Sun, A. P. Petropulu, and H. V. Poor, "MIMO radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 98–117, 2020.

[4] K. V. Mishra, M. R. Bhavani Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter wave joint radar-communications: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 100–114, 2019.

[5] S. H. Dokhanchi, B. S. Mysore, K. V. Mishra, and B. Ottersten, "A mmWave automotive joint radar-communications system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 3, pp. 1241–1260, 2019.

[6] G. Duggal, S. Vishwakarma, K. V. Mishra, and S. S. Ram, "Doppler-resilient 802.11ad-based ultrashort range automotive joint radar-communications system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 5, pp. 4035–4048, 2020.

[7] S. Alland, W. Stark, M. Ali, and A. Hedge, "Interference in automotive radar systems: Characteristics, mitigation techniques, and future research," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 45–59, 2019.

[8] Z. Slavik and K. V. Mishra, "Cognitive interference mitigation in automotive radars," in *IEEE Radar Conference*, 2019, pp. 1–6.

[9] S. Na, K. V. Mishra, Y. Liu, Y. C. Eldar, and X. Wang, "TenD-SuR: Tensor-based 4D sub-Nyquist radar," *IEEE Signal Processing Letters*, vol. 26, no. 2, pp. 237–241, 2018.

[10] K. V. Mishra and Y. C. Eldar, "Sub-Nyquist radar: Principles and prototypes," in *Compressed Sensing in Radar Signal Processing*, A. D. Maio, Y. C. Eldar, and A. Haimovich, Eds. Cambridge University Press, 2019, in press.

[11] K. V. Mishra, S. Mulleti, and Y. C. Eldar, "RaSSteR: Random sparse step-frequency radar," *arXiv preprint arXiv:2004.05720*, 2020.

[12] M. W. Maier, "Non-uniform PRI pulse-Doppler radar," in *Southeastern Symposium on System Theory*, Tuscaloosa, AL, Mar. 7-9, 1993, pp. 164–168.

[13] J. Li and Z. Chen, "Research on random PRI PD radar target velocity estimate based on NUFFT," in *IEEE CIE International Conference on Radar*, 2011, pp. 1801–1803.

[14] W. P. Plessis, "Simultaneous unambiguous range and Doppler through non-uniform sampling," in *Proc. IEEE Radar Conf.*, Florence, Italy, Sept. 21-25, 2020.

[15] S. Sedighi, B. Shankar, K. V. Mishra, and B. Ottersten, "Optimum design for sparse FDA-MIMO automotive radar," in *Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 913–918.

[16] K. V. Mishra, I. Kahane, A. Kaufmann, and Y. C. Eldar, "High spatial resolution radar using thinned arrays," in *IEEE Radar Conference*, 2017, pp. 1119–1124.

[17] C.-Y. Chen and P. P. Vaidyanathan, "Minimum redundancy MIMO radars," in *IEEE International Symposium on Circuits and Systems*, 2008, pp. 45–48.

[18] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, 2010.

[19] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, 2011.

[20] S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1377–1390, 2015.

[21] M. Wang and A. Nehorai, "Coarrays, MUSIC, and the Cramér-Rao bound," *IEEE Transactions on Signal Processing*, vol. 65, no. 4, pp. 933–946, 2017.

[22] C.-L. Liu and P. P. Vaidyanathan, "Super nested arrays: Linear sparse arrays with reduced mutual coupling - Part I: Fundamentals," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3997–4012, 2016.

[23] C.-L. Liu and P. Vaidyanathan, "Super nested arrays: Linear sparse arrays with reduced mutual coupling—Part II: High-order extensions," *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4203–4217, 2016.

[24] H. Qiao and P. Pal, "Guaranteed localization of more sources than sensors with finite snapshots in multiple measurement vector models using difference co-arrays," *IEEE Transactions on Signal Processing*, vol. 67, no. 22, pp. 5715–5729, 2019.

[25] S. Qin, Y. D. Zhang, M. G. Amin, and A. M. Zoubir, "Generalized coprime sampling of Toeplitz matrices for spectrum estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 1, pp. 81–94, 2017.

[26] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, "The rise of radar for autonomous vehicles: Signal processing solutions and future research directions," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 20–31, 2019.

[27] A. Gupta, U. Madhow, and A. Arbabian, "Super-resolution in position and velocity estimation for short-range MM-wave radar," in *Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 1144–1148.

[28] M. M. Hyder and K. Mahata, "Range-Doppler imaging via sparse representation," in *IEEE Radar Conference*, 2011, pp. 486–491.

[29] E. J. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

[30] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.

LEADER-FOLLOWER MULTI-AGENT SYSTEMS: A MODEL PREDICTIVE CONTROL SCHEME AGAINST COVERT ATTACKS

Francesco Tedesco, Domenico Famularo and Giuseppe Franzè

Università della Calabria, Via Pietro Bucci, Cubo 42-C, Rende (CS), 87036, ITALY,
{francesco.tedesco,domenico.famularo,giuseppe.franze}@unical.it

ABSTRACT

In this paper, a resilient distributed control scheme against covert attacks for constrained multi-agent networked systems is developed. The idea consists in an adequate deployment of predictive arguments with a twofold aim: detection of malicious agent behaviors and control actions implementation to mitigate as much as possible undesirable knock-on effects.

Index Terms— Resilient control, Multi-agent systems, Covert Attacks, Constraints

1. INTRODUCTION

The integration process of heterogeneous devices aiming at the regulation of physical processes via communication networks consists in providing such systems with better operational and management capabilities, as well as reducing its costs, i.e. the Cyber-Physical Systems (CPSs) paradigm [2]. Within this framework, attention will be devoted to address the attack detection problem by focusing on a class of intelligent coordinated intrusions known as covert attacks (see [15], [8], [13] and references therein). It is worth to underline that such type of adverse phenomena become severe (and interesting) within multi-agent system configurations, see e.g. [14], [7], and can be successfully managed via a switching control strategy by borrowing distributed model predictive control (MPC) ideas [3].

To this end, for each agent a moving target plant behavior is imposed via a control policy whose current action is obtained by randomly choosing among three admissible different and compatible each other control sequences. Two different paradigms are proposed: the first scheme, hereafter denoted as **Delay-MPC**, is a dual-mode receding horizon MPC controller making use of set-theoretic arguments [5], while the second one, namely **N-MPC**, is a robust $N-$steps MPC scheme which is *ad-hoc* tuned to enhance the features of the **Delay-MPC** controller [6]. Hence, two **N-MPC** sequences are off-line computed and stored in the actuator buffer, while the third control action on-line comes out from the **Delay-MPC** algorithm. During the time interval defined by the prediction horizon length N, the **N-MPC** sequences are randomly combined according to a cryptographically secure pseudo-random number generator whose initial seed is shared with the detector on the controller side. Then, the off-line MPC sequences are updated when the time interval tuned w.r.t N is elapsed, a communication channel refresh is performed according to the software rejuvenation approach [11], [9], because on the controller side it is unknown when an attack occurs. By following this *modus operandi*, the covert attack cannot remain indefinitely stealthy and can be detected by exploiting one-step controllability results. The possibly dangerous effects are mitigated by exclusively resorting to the command inputs stored in the actuator buffer until the periodic communication medium refresh takes place.

2. PROBLEM FORMULATION

In the sequel, the class of grid leader-follower configurations with a unique path connecting each follower to a leader is considered.

The multi-agent system consisting of L subsystems is organized with respect to the grid topology via the operator $level(i) : \{1,\dots,L\} \to \mathbb{Z}_+$, which provides the position of the $i-th$ agent along the grid. Then, r levels result each one collecting l_i agents, with any sub-system denoted by Σ_j^i, and j accounting for the $level(\cdot)$. Therefore:

$$\mathbf{\Sigma_j^i}: \begin{cases} x_j^i(t+1) &= A_j^i x(t) + B_j^i u_j^i(t) + d(t) \\ y_j^i(t) &= x_j^i(t) + n(t), \\ & i = 1,\dots,r,\ j = 1,\dots,l_i \end{cases} \tag{1}$$

where $t \in \mathbb{Z}_+ := \{0,1,\dots\}$, $x(t) \in \mathbb{R}^n$ denotes the state, $u(t) \in \mathbb{R}^m$ the input, $y(t) \in \mathbb{R}^n$ the set of state measurements, $d(t) \in \mathcal{D} := \{d \in \mathbb{R}^n \,|d^T d \le \bar{d}\} \subset \mathbb{R}^n$, $\forall t \in \mathbb{Z}_+$, and $n(t) \in \mathcal{N} := \{n \in \mathbb{R}^n \,|n^T n \le \bar{n}\} \subset \mathbb{R}^n$, $\forall t \in \mathbb{Z}_+$, the process and measurement noises, respectively. Without loss of generality the sets $\mathcal{D}$ and $\mathcal{N}$ are assumed to be identical for all the agents Σ_j^i, $I = 1,\dots,r$, $j = 1,\dots,l_i$. Moreover, (1) is subject to input and state constraints:

$$\begin{aligned} u_j^i(t) \in \mathcal{U}_j^i &:= \{u_j^i \in \mathbb{R}^m : {u_j^i}^T u_j^i \le (u_j^i)_{max}\}, \\ x_j^i(t) \in \mathcal{X}_j^i &:= \{x_j^i \in \mathbb{R}^n : {x_j^i}^T x_j^i \le (x_j^i)_{max}\}, \forall t \end{aligned} \tag{2}$$

The following definitions are used: *set of neighbors*: $\mathcal{N}_j^i := \{q \in \{1,\dots,i-1,i+1,\dots,L\} : level(q) \equiv level(i)\}$; *father*: the operator $(pre(i), pre(j)) : \{1,\dots,r\} \times \{1 \dots, l_i\} \to$

$\{1,\dots,r\}\times\{1\ \dots,l_i\}$ identifies the father of Σ_j^i; *set of children*: $\mathcal{H}_j^i := \{h \in \{1,\dots,l_{j+1}\} : \Sigma_j^i \text{ is a father of } \Sigma_{j+1}^h\}$ Moreover, it is assumed that each Σ_j^i sends to neighbors and children a packet containing its state trajectory $\mathbf{x}_j^i(\cdot)$.

An adversary launches the following covert attacks to the integrity of control actions and sensor measurements [15]:
actuation channel tampering: $u_j^i(t) = w_j^i(t) + \Delta w_j^i(t), \forall t$, where $\Delta w_j^i(t) \in \mathbb{R}^m$ is unknown and not bounded;
$\Delta z_j^i(t)$ **affects computation on the plant output measurements and subsequent subtraction from the measured plant output**: $z_j^i(t) = y_j^i(t) - \Delta z_j^i(t),\ \forall t$.
Here, the following problem is addressed:
Given the multi-agent system (1) subject to covert attacks, determine a distributed state-feedback resilient control policy as

$$\begin{aligned} u_j^1(t) &= g(x_j^1(t), \{x_k^1(t)\}),\ \forall k \in \mathcal{N}_j^1, j = 1,\dots,l_1, \\ u_j^i(t) &= g(x_j^i(t), x_{pre(j)}^{pre(i)}(t), \{x_k^i(t)\}),\ \forall\, k \in \mathcal{N}_j^i, \\ &\quad i = 2,\dots,r; j = 1,\dots,l_i, \end{aligned} \tag{3}$$

compatible with (2) such that, starting from any admissible initial condition $[x^{1^T}(0),\ x^{2^T}(0)\dots x^{L^T}(0)]^T$, *the grid topology* $\mathcal{G}$ *is indefinitely kept despite any admissible disturbance/noise realization and covert attack occurrence.*

3. RESILIENT CONTROL STRATEGY

The following aspects will be relevant: 1) no *a-priori* information are available on the attack occurrence; 2) any delivered command $w_j^i(\cdot)$ could be affected by anomalies introduced by the adversary; 3) any control action will be exclusively computed on the remote side.

Starting from these preliminaries the controller will be designed by exploiting MPC ideas that well adapt to *worst-case* scenarios. Since the *controller-to-actuator* link is unreliable, at each time instant any received data packet $w_j^i(\cdot)$ could be affected by adversary malicious actions and, once the attack has been detected, the resulting control architecture has to operate in an open-loop fashion. To this end, an **Actuator Buffer** unit is introduced to store feasible command input sequences to be used during the on-line operations. On the other hand, to update the control action $w_j^i(\cdot)$ a reliable state condition is required. Then, a **Controller Buffer** is used to take care of the last received and attack-free state measurement $z_j^i(\cdot)$. As the covert attack is concerned, the aim is to reduce the level of plant/controller knowledge pertaining to the covert agent side in order to mitigate its effect. Therefore, starting from the detection time instant, the resilient control strategy must be capable to ensure a safe open-loop behavior for a possible long time period, i.e. the control horizon length N of the MPC controller must be maximized compatibly with the unavoidable computational demands. Then, the resilient scheme of Fig. 1 is hereafter considered.

Initially, two admissible input sequences $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I}$ and $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II}$ compatible with all the prescribed constraints (2) are off-line designed and stored in the **Actuator Buffer**.

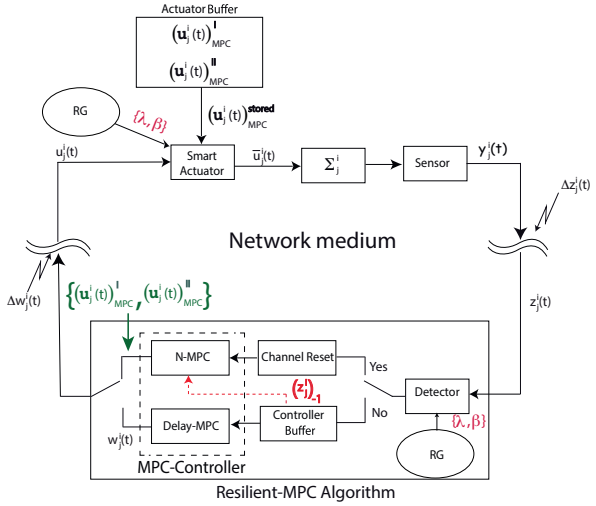


Fig. 1. The resilient control architecture against covert attacks

During the on-line phase the **Detector** unit is in charge to establish if the agent Σ_j^i is under attack. Starting from the generic initial time instant t, if $N-1$ time instants are elapsed ($t \leftarrow t+N-1$) (label **Yes**), a **Channel Reset** procedure starts by instantaneously disconnecting the network medium and taking the following actions: 1) the **Attack mode** is activated: send/receive operations are interrupted and, according to the grid topology $\mathcal{G}$, all the agents belonging to the upper levels are instantaneously disconnected from the associated local controllers; 2) on the agent side and for all agents operating in **Attack mode** , the open-loop dynamical behavior of Σ_j^i is exclusively regulated with the stored moves $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{stored}$; 3) on the controller side, novel admissible sequences $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I}$ and $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II}$ are computed by using the received data $z_j^i(t)$; 4) as soon as such sequences are obtained, the channel is restored with to the controller-to-actuator channel being totally secure for a single time instant, then $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I}$ and $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II}$ are transmitted and stored in the **Actuator Buffer**.
Conversely (label **No**) the new command $w_j^i(t)$ is on-line computed and sent to the **Smart Actuator**.
The **Smart Actuator** is in charge to select the current command input $\bar{u}_j^i(t)$ by exploiting the following data: the received command $u_j^i(t)$, the stored inputs $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I}$ and $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II}$ and two scalars $\lambda \in [0,\ 1] \subset \mathbb{R}$ and $\beta \in \{0,1\} \subset \mathbb{Z}_+$ generated by a cryptographically secure pseudo-random number generator unit **RG**. Notice that the same seed number is shared between the smart actuator and the detector. These scalars are exploited as follows: λ is used to define the following convex combination:

$$\left(\mathbf{u}_j^i(t)\right)_{MPC}^{stored} := \lambda(t)\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I} + (1-\lambda(t))\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II} \tag{4}$$

β accounts for the input that will be selected: if $\beta = 1$

the input $u_j^i(t)$ is used, otherwise the stored command $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{stored}$ is selected.
For the sake of simplicity and without loss of generality scalars β and λ are free of subscript and apex. Notice that a *twin* identical seed pseudo-random number generator is associated to the **Detector** unit and correctly synchronized to the agent side. Hence, the resulting output $y_j^i(t)$ is transmitted and the **Detector** will evaluate $z_j^i(t)$: if the response is **Yes** the communication medium is immediately disconnected from the open-loop chain, otherwise an attack-free scenario is considered. Finally it is important to underline that, since the **Channel Reset** procedure periodically takes place at each N time instants (the length of the MPC sequences), there necessary exists an admissible control law capable to keep the regulated state trajectory within a guaranteed region until the new sequences $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I}$ and $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II}$ are received.

4. THE DISTRIBUTED MPC CONTROLLER

The above reasoning is focusing the following issues: **a)** the stored sequences $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{I}$ and $\left(\mathbf{u}_j^i(t)\right)_{MPC}^{II}$ have to be admissible and compatible each other; **b)** the computation of the MPC sequences requires that each unit Σ_j^i, whenever necessary, could safely proceed in an open-loop fashion until this task is accomplished. The **MPC-Controller** consists of two coupled MPC units: the first (hereafter named **Delay-MPC**) is determined under the constraint that each resulting command could be successively usable for the next time instant, whereas the second MPC-based unit (**N-MPC**) is designed by choosing as initial conditions those provided by the **Delay-MPC** and with the control horizon length N as large as possible. These controllers are designed by assuming that the multi-agent system operates under attack-free conditions: $z_j^i(t) \equiv y_j^i(t)$ and $w_j^i(t) \equiv u_j^i(t), \forall i, j, \forall t$, and by following *mutatis mutandis* the technicalities reported in [4] and [7]. The control action resulting from the unit **Delay-MPC** is based on the computation of sequence of **Predecessor Families** $\{\Xi_j^i\}_k, i = 1, \ldots, r; j = 1, \ldots, l_i$; while the second controller provides the following sequences of N control moves

$$(\mathbf{u_j^i})_{MPC}^{I} =: \{(u_j^i)_k^I(t)\}_{k=0}^{N-1}, (\mathbf{u_j^i})_{MPC}^{II} =: \{(u_j^i)_k^{II}(t)\}_{k=0}^{N-1}$$

where the last entries , $(u_j^i)_{N-1}^I$ and $(u_j^i)_{N-1}^{II}$ can be consecutively applied $k_o^{\mathcal{G}}$ times.

5. DETECTOR

For each agent Σ_j^i, this unit is in charge to verify the admissibility of transmitted data $z_j^i(t)$. By resorting to the arguments developed in [12], the idea consists on using the concept of expected one-step prediction set $(\mathbf{Z}_j^i)^+$. To this end, the **Controller Buffer** is equipped with a counter $(Count_j^i)_{contr}$ initialized at $N-1$ that decreases by one at each time instant, whereas the *twin* **RG**, synchronized with the pseudo-random generator on the plant side, provides the current pair $(\lambda(t), \beta(t))$. Then:

$$\begin{array}{l}(\mathbf{X}_j^i)^+\left(z_j^i(t),(u_j^i)_{curr}\right) := \{(z_j^i)^+ \in \mathbb{R}^n: \\ (z_j^i)^+ = A_j^i z_j^i(t) + B_j^i \bar{u}_j^i(t) + (B_j^i)_d d, \forall\, d \in \mathcal{D}\} \subset (\Xi_j^i)_{s-1}\end{array} \quad (5)$$

where $z_j^i(t) \in (\Xi_j^i)_s$ is the available information at the previous time instant,

$$\bar{u}_j^i(t) = \begin{cases} (u_j^i)^{-1}(t), & \text{if } \beta(t) = 1 \\ (\mathbf{u}_j^i)_{MPC}^{stored_k} & \text{otherwise} \end{cases}$$

and $(\mathbf{u}_j^i)_{MPC}^{(stored)_k}$, with $k = N - 1 - (Count_j^i)_{contr}$, is the convex combination by $\lambda(t)$ of the $k-th$ entries of $(\mathbf{u}_j^i)_{MPC}^I(t)$ and $(\mathbf{u}_j^i)_{MPC}^{II}(t)$. Then, one has that:

$$\begin{array}{l}(\mathbf{D}_j^i)^+(z_j^i(t+1)) := \\ \begin{cases} \text{attack, if } z_j^i(t+1) \notin (\mathbf{X}_j^i)^+\left(z_j^i(t), (\bar{u}_j^i)(t)\right) \\ \text{no attack, otherwise} \end{cases}\end{array} \quad (6)$$

6. CHANNEL RESET CONTROL ACTIONS

By referring to the scheme of Fig. 1, this unit is periodically active (**Yes** label) after every $N - 1 + k_o^{\mathcal{G}}$ time instants. In particular, the following actions are performed: 1) the communication channels are instantaneously shut down; 2) the plant proceeds in an open-loop fashion under the action of the admissible control input $(\bar{u}_j^i)_{N-1}$ stored in the **Actuator Buffer**; 3) new admissible sequences $(\mathbf{u}_j^i)_{MPC}^I(t)$ and $(\mathbf{u}_j^i)_{MPC}^{II}(t)$ are computed by using the attack-free measurement $(z_j^i)_{-1}$ stored in the **Controller Buffer** as the initial state condition.
Then, the *modus operandi* of the reset process can be summarized as follow. At each time instant t, the **Detector** evaluates the received measurement $z_j^i(t)$. If an attack is detected at a certain time instant $\bar{t} \in [t,\ t+N-1]$, then the **Attack Mode** is activated, an instantaneous and complete disconnection of the network medium (controller-to-actuator and sensor-to-controller channels) takes place and the agent proceeds in an open-loop by using the remaining number of stored commands $(t + N - 1 + k_o^{\mathcal{G}}) - \bar{t}$. Since **Predecessor Families** are computed explicitly taking care of neighbors and father agents along the grid $\mathcal{G}$, then the **Attack Mode** must be extended to neighbors and the upper levels, i.e. plant and controller are disconnected each other for all the pairs $(\Sigma_s^i, C_s^i), \forall s \in \mathcal{N}_j^i$ and (Σ_q^p, C_q^p), $p = 1, \ldots, i-1,\ q = 1, \ldots, l_p$.

On the remote side, the attack detection (**No** label) imposes the construction of a new admissible input sequences $(\mathbf{u}_j^i)_{MPC}^I(t)$ and $(\mathbf{u}_j^i)_{MPC}^{II}(t)$ by following properties and prescriptions outlined in the previous section. To this end, the remote side must be accorded to the plant dynamical evolution at the time instant $t+N-1+k_o^{\mathcal{G}}$, this leads to the computation of the state evolution starting from the event $(\bar{t}, (z_j^i)_{-1})$ under the action of $(\mathbf{u}_j^i)_{MPC}^{stored}$. Once the new input sequences have been computed, the **Actuator Buffer** can be updated by re-activating the communication network. To comply with

this reasoning, the following assumptions are required, see [12]: *a guaranteed attack-free communication between the controller and the plant and vice-versa can be re-established in at most $k_o^{\mathcal{G}}$ time steps; for at least a time instant after the network recovery, the controller-to-actuator channel is guaranteed to be attack-free.*

Finally, the **Smart Actuator** logic is:

$$\bar{u}_j^i(t) = \begin{cases} (u_j^i)^{-1}(t), & \text{if } \beta(t)=1 \\ (\mathbf{u}_j^i)_{MPC}^{stored_k} & \text{if } \beta(t)=0 \\ (\mathbf{u}_j^i)_{MPC}^{stored_k} & \textbf{Attack Mode}\text{==true} \end{cases} \quad (7)$$

7. SIMULATIONS

In this section, simulations show the effectiveness of the proposed **Resilient MPC** algorithm (referred as *resilient*) by investigating countermeasure features and contrasting with a *no-resilient* competitor being a **Resilient MPC** scheme only equipped with a **Delay-MPC** block and not endowed with detection capabilities.

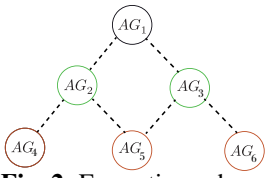


Fig. 2. Formation scheme

Consider six robots organized as depicted in Fig. 2 and described by

$$x(t+1) = \Phi x(t) + Gu(t)$$

$$\Phi = \begin{bmatrix} I_2 & \Delta t\, I_2 \\ 0_2 & I_2 \end{bmatrix},\ G = \begin{bmatrix} \frac{(\Delta t)^2\, I_2}{2} \\ \Delta t I_2 \end{bmatrix},\ \Delta t = 0.1\ sec.$$

where the state accounts for position (x_1, x_2) and velocity (x_3, x_4) components, while the input $u \in \mathbb{R}^2$ for the acceleration vector (m/s^2). According to the multi-agent system description (1), one has: $AG_1 \leftarrow \Sigma_1^1$, $AG_2 \leftarrow \Sigma_1^2$, $AG_3 \leftarrow \Sigma_2^2$, $AG_4 \leftarrow \Sigma_1^3$ $AG_5 \leftarrow \Sigma_2^3$ $AG_6 \leftarrow \Sigma_3^3$. Moreover the following constraints are prescribed:

$$|(u_j^i)_k(t)| \leq 1[m/s^2], i=1,2,3; j=1,\dots l_i; k=1,2; \quad (8)$$

$$\begin{aligned} &h \leq (x_j^i)_1 - (x_q^p)_1 \leq (3/2)h, \\ &\forall (i,j,p,q) \in \{(1,1,2,1),(2,2,1,1),(2,1,3,1), \\ &(3,2,2,1),(2,2,3,2),(3,3,2,2)\} \end{aligned} \quad (9)$$

$$\begin{aligned} &h \leq (x_j^i)_2 - (x_q^p)_2 \leq (3/2)h, \\ &\forall (i,j,p,q) \in \{(1,1,2,1),(1,1,2,2),(2,1,3,1), \\ &(2,1,3,2),(2,2,3,2),(2,2,3,3)\},\ h = 0.1m. \end{aligned} \quad (10)$$

In the sequel, the following operating scenario is considered: *It is required that the robot team leader Σ_1^1 is driven to the origin while keeping the formation constraints despite the occurrence of two covert attacks during the time intervals* $[4.9,\ 6.9]$ *s and* $[22.9,\ 26.9]$ *s on the agents Σ_1^1 and Σ_2^3 respectively.*

Numerical results are collected in Figs. 3-5. At $t = 4.9\,s$ when the attack on Σ_1^1 takes place, the **RDC** *no resilient* algorithm applies the computed command (see Fig. 3) without recognizing that the resulting state trajectory $x_1^1(\cdot)$ is going to *overcome* prescribed constraints, see Figs. 4 (dot-dashed black line) and 5.b (solid black line). Moreover despite of the end of the attack at $t = 6.9\,s$, the state measurement is conveyed to Σ_1^1 that in turn is no longer able to compute a feasible command: the default input $u_1^1(t) = [0,0]^T$ is applied. Similar comments can be done when the attack on the agent Σ_1^3. occurs. Finally, notice that formation requirements (9)-(10) are always satisfied, see Fig. 5.

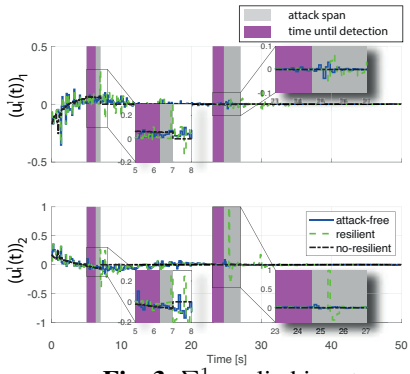


Fig. 3. Σ_1^1 applied inputs

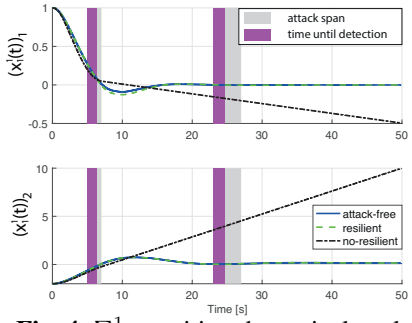


Fig. 4. Σ_1^1 : position dynamical evolution

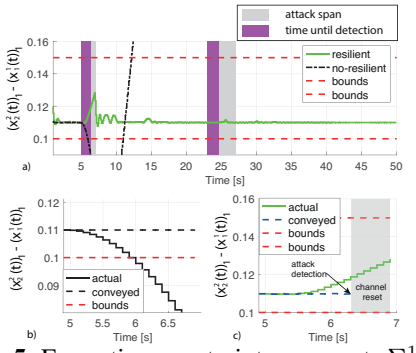


Fig. 5. Formation constraints on agents Σ_1^1 and Σ_2^2

8. REFERENCES

[1] F. Blanchini, and S. Miani, "Set-theoretic methods in control", *IEEE Boston: Birkhäuser*, 2008.

[2] A. Burg, A. Chattopadhyay, and K. Lam, "Wireless Communication and Security Issues for Cyber–Physical Systems and the Internet-of-Things", *Proceedings of the IEEE*, vol. 106(1), pp. 38-60, 2018.

[3] P. D. Christofides, R. Scattolini, D. M. de la Pena, and J. Liu, "Distributed model predictive control: A tutorial review and future research directions", *Computers & Chemical Engineering* vol., 51, pp. 21-41, 2013.

[4] D. Famularo, G. Franzè, W. Lucia, and C. Manna , "A reconfiguration control framework for constrained systems with sensor stuck faults", *International Journal of Robust and Nonlinear Control*, vol. 29(4), pp. 1150-1164, June 2019.

[5] G. Franzè, F. Tedesco, and D. Famularo, "Model predictive control for constrained networked systems subject to data losses", *Automatica*, vol. 54, pp. 272-278, 2015.

[6] G. Franzè, F. Tedesco, and W. Lucia, "Resilient control for cyber-physical systems subject to replay attacks", *IEEE Control Systems Letters*, vol. 3-4, pp. 984-989, 2019.

[7] G. Franzè, F. Tedesco, and D. Famularo, "A distributed resilient control strategy for leader-follower systems under replay attacks", *In 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT)*, IEEE, 2020, vol. I, pp. 469-474.

[8] Ghaderi M, Gheitasi K and Lucia W (2020). A Blended Active Detection Strategy for False Data Injection Attacks in Cyber-Physical Systems. *IEEE Transactions on Control of Network Systems*, 2020.

[9] P. Griffioen, R. Romagnoli, B. H. Krogh, and B. Sinopoli B (2019, December) Secure Networked Control via Software Rejuvenation", *58th Conference on Decision and Control (CDC)*, IEEE, 2019, pp. 3878-3884.

[10] A. Hoehn, and P. Zhang, "Detection of covert attacks and zero dynamics attacks in cyber-physical systems", *In 2016 IEEE American Control Conference (ACC)*, IEEE, 2016, pp. 302-307.

[11] Y. Huang, C. Kintala, N. Kolettis, and N. D. Fulton, "Software rejuvenation: Analysis, module and applications", *In Twenty-fifth international symposium on fault-tolerant computing*, 1995, pp. 381-390.

[12] W. Lucia, B. Sinopoli, and G. Franzè, "A set-theoretic approach for secure and resilient control of cyber-physical systems subject to false data injection attacks", *In 2016 Science of Security for Cyber-Physical Systems Workshop (SOSCYPS)*, IEEE, 2016, pp.1-5.

[13] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding schemes for securing cyber-physical systems against stealthy data injection attacks", *IEEE Transactions on Control of Network Systems*, vol. 4(1), pp. 106-117, 2016.

[14] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach", *IEEE Transactions on Automatic Control*, vol. 57(1), pp. 90-104, 2011.

[15] R. S. Smith, "Covert misappropriation of networked control systems: Presenting a feedback structure", *IEEE Control Systems Magazine*, vol. 35(1), pp.82-92, 2015.

STATE-OF-THE-ART AND DIRECTIONS FOR THE CONCEPTUAL DESIGN OF SAFETY-CRITICAL UNMANNED AND AUTONOMOUS AERIAL VEHICLES

Saad Bin Nazarudeen, Jonathan Liscouët

Concordia University, Gina Cody School of Engineering and Computer Science, Department of Mechanical, Industrial and Aerospace Engineering, Montreal, Canada.

ABSTRACT

Unmanned and Autonomous Aerial Vehicles (UAV/AAV) must be safe and reliable to prevent catastrophic accidents in population-dense areas. The study reveals the absence of a comprehensive UAV/AAV design for reliability approach in the open literature; in particular, there is no conceptual design methodology including safety and reliability considerations in the sizing. This finding leads to investigating the relevance of pursuing this research direction and identifying the challenges to address. For this matter, a straightforward approach combining sizing, systematic redundancy, controllability, and reliability assessments compares a conventional to a redundant design in a case study. The reliability analysis confirms that the redundant design is fault-tolerant and potentially highly reliable. However, the total mass almost doubles due to the lack of sizing and redundancy optimization. Plus, there is a high risk of under-sizing due to the limitations of a straightforward approach. This result emphasizes the need to develop a new conceptual design methodology based on sizing, including safety and reliability considerations. The paper concludes with research directions towards this goal. Thus, optimized redundant designs will contribute to the emergence of UAV/AAV for safety-critical applications in the near future.

Index Terms — Unmanned Aerial Vehicle (UAV), Autonomous Aerial Vehicle (AAV), Reliability, Redundancy, Fault-tolerant, Multicopter, Conceptual Design, Sizing.

1. INTRODUCTION

Unmanned and Autonomous Aerial Vehicles (UAV and AAV) were under active development over the past sixteen years [1]. They have shown promising applications in urban surveillance, agriculture, media coverage, logistics, deliveries, flying taxis, or flying ambulances that would change our daily lives. Emerging autonomous control has evolved UAV into AAV [2]–[4]. These vehicles can accomplish a preassigned task without human operators, especially high-risk operations like air taxiing, painting skyscrapers, or aerial firefighting. Such operations are safety-critical. Failure leading to an uncontrollable UAV can cause a catastrophic accident if it collides with humans, aircraft, helicopters, or infrastructures. This damage is increasing with its mass and size. Therefore, designs with high safety and reliability are required.

The primary objectives of the conceptual design are to find design concepts, evaluate them and select promising ones. The focus here is on the evaluation. The state-of-the-art conceptual design of UAV/AAV focuses on propulsor configurations and evaluates them through sizing for performance (e.g., overall weight, payload, range, maximum speed) [5]–[9]. Additionally, some studies evaluate the reliability of propulsor configurations based on controllability [10], [11]. Others optimize control allocation to maximize the flight reliability of a given configuration [12], [13]. Some works relate to the safety of UAVs by introducing fault-tolerant control strategies [14] which can be classified into passive [15], [16], and active [17], [18]. Several works produce unique designs and design features to increase resilience [19], [20]. However, control strategies and unique design solutions are out-of-scope of a conceptual design methodology. It stands out that none of the works mentioned above include safety and reliability considerations in the sizing.

This finding leads to investigate the relevance and challenges of a conceptual design with safety and reliability considerations. For the matter, a straightforward approach combining sizing, systematic redundancy, controllability, and reliability assessments is applied to a case study. The case study compares a conventional to a redundant design in terms of total weight and reliability. It leads to the limitations of a straightforward approach and the challenges of a redundant design and thus provides research directions for an effective conceptual design methodology.

The paper is structured as follows. Section 2 describes the case study comparing a conventional with a redundant design followed by conclusions and research directions in section 3.

2. CASE STUDY

In this section, a case study compares a conventional with a redundant UAV using criteria like fault-tolerance, reliability,

and mass to identify the challenges posed by redundant design.

2.1. Medium-sized midrange multicopter

The selected case study represents medium-sized media drones, mail delivery drones, and surveillance drones that usually operate over heavily populated areas.

This design does not represent the typical size of safety-critical applications like air taxis, flying ambulances, and aerial firefighting. The choice of a medium-sized drone is driven by the limitation of the selected sizing tool (see Section 2.2.), which is designed to handle a maximum payload of 10 kg only.

2.2. Sizing

Since no methodology in the open literature addresses safety or reliability considerations in the sizing, a straightforward approach has been applied to incorporate redundancy. The tool named Flyeval [21] is used to size a conventional architecture. It doesn't have any means to apply redundancy; therefore, parallel components are added manually. The weight of each redundant component is added to the total weight, looping back to the performance evaluation and eventually to another sizing loop until converging. This approach avoids developing a sizing tool, but it does not optimize the redundancies and neglects the effect of failure cases on sizing.

Flyeval is proposed in academic research for validating drone sizing methodologies [22], [23]. It allows the user to perform sizing according to preliminary design requirements and select off-the-shelf components from a database. It then generates a mathematical model and calculates the resulting key performance criteria (forward speed, flying range, and hovering time) so that the user can efficiently verify or refine its component selection.

In the case study, the primary mission requirement considered for the quadcopter sizing is to carry a payload of 10 kg within a 5 km range.

2.3. Reliability analysis

The typical first rule of safety-critical design standards is: A catastrophic failure condition must not result from a single failure and must be extremely improbable [24, Pt. VTOL.2510(a)] . A catastrophic failure condition is defined as: High impact crash is imminent and unavoidable with the vehicle's destruction. Severe injuries or the death of people on the ground is possible. Infrastructures can be damaged heavily. Here, a catastrophic failure condition shall have a probability of occurrence less than or equal to 10^{-7} [24], [25].

Other rules for less critical conditions (hazardous, major, and minor) also apply, but the presented reliability analysis focuses on evaluating the conventional and redundant designs against this first rule only for simplicity and conciseness.

Some studies on fault-tolerant control of quadcopters have proposed emergency landing procedures [26]–[28] to avoid a single rotor failure to lead to a catastrophic condition. It is a significant improvement for the reliability of the quadcopter. However, these procedures do not maintain the yaw control. In the studied case, the failure can occur in cruise flight and the control scheme must completely stop the speeding vehicle and engage a hover mode before any controlled descent. This procedure requires sufficient control of all control axis. That is why the present reliability analysis assumes that a catastrophic failure condition can be avoided only if the UAV keeps full or degraded control of all control axis.

The reliability analysis of the conventional and redundant designs is performed with standard FMEA and RBD methodologies [29] and boolean algebra. The reliability computation is based on independent and random failures, leading to constant failure rates and the convenient exponential distribution. The probability of failure is expressed as:

$$F(t) = 1 - e^{-\lambda t} \qquad (1)$$

Where, λ is the failure rate (h^{-1}) and t is the exposure time (h).

The exposure time corresponds to the maximum flight time of 22 minutes between two charges (see Section 2.4.) and is enforced by the assumption of a built-in-test procedure confirming the functionality of each component at power-up. The failure rates summarized in Figure 1 are based on the orders of magnitude of modern high-end transport category aircraft equipment.

2.4. Conventional design

The conventional design of the quadcopter consists of frame, battery, power distribution board, flight controller, onboard instruments, and four sets of electronic speed controllers, motors, and propellers with no redundancy.

2.4.1. Conventional design reliability analysis

Failure Mode and Effect Analysis (FMEA) is a simple and widely used reliability analysis that considers the effects of each component's single failures and assesses the failure severity [29]. It shows that each component of the conventional quadcopter is susceptible to at least one catastrophic effect. Hence, it violates the typical first rule of safety-critical design standards. It can be concluded that the conventional design is unfit for any safety-critical application.

2.5. Redundant design

The redundant design is obtained by incorporating redundancies systematically into the conventional design: (1) A redundant battery and a battery management system are added;

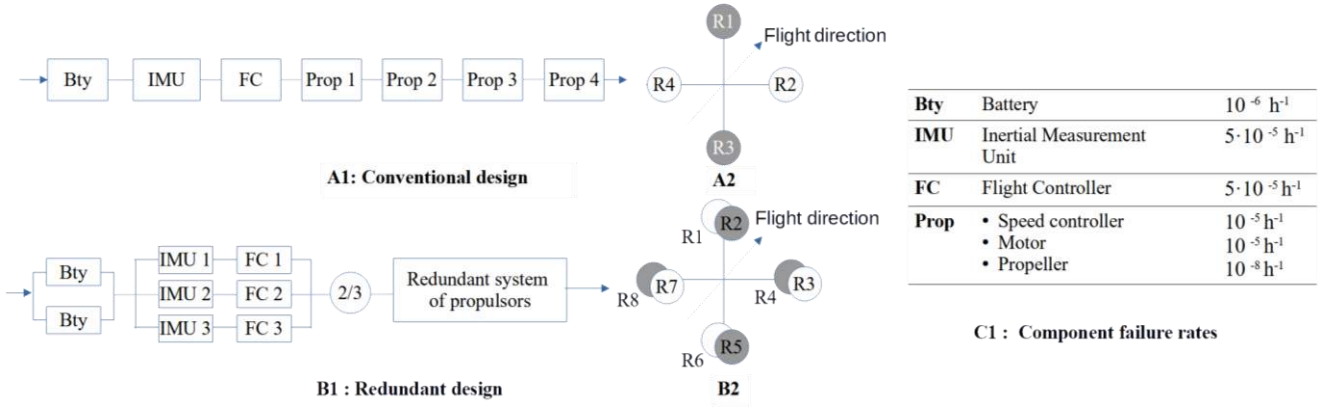


Figure 1. RBD of conventional design (A1) and redundant design (B1), corresponding configurations (A2 and B2) and component failure rates (C1)

(2) Onboard instruments and the flight controller are tripled and managed with a majority voting algorithm. In case of failure, the service will continue with the two agreeing flight controllers; (3) Quadcopter configuration is modified into a co-axial configuration for redundancy and allowing developing control failure mitigating strategies; (4) As the design is modified, a new off-the-shelf frame is selected considering the increase in maximum take of weight due to redundant components. The thickness of the plate of the frame is increased for rigidity. The arm and motor mass are increased to reflect dualization.

2.5.1. Redundant design reliability analysis

Unlike conventional design, the redundant design is robust to single failures. Therefore, the reliability analysis needs to focus on the combination of failures. Fault Tree Analysis (FTA) and Reliability Block Diagram (RBD) analysis [29] are more suited than an FMEA for this purpose. RBD is selected because it has the advantage of providing a physical illustration of the drone architecture.

For comparison, an RBD is constructed for both conventional and redundant design according to the ability of the design to perform the emergency landing procedure to avoid a catastrophic condition. The reliability of the majority voting redundancies is obtained from a k-out-of-n redundancy calculation as follows:

$$R_{2/3}(t) = 3 \cdot R_i(t)^2 \cdot \big(1 - R_i(t)\big) + R_i(t)^3 \qquad (2)$$

Where, $R_i(t)$ is the reliability of the identical components of the redundancy, in effect, each assembly of an inertial measurement unit and flight controller.

An evaluation of reliability based on controllability is presented in [11]. This approach applies here to the propulsor system. A propulsor is the assembly of a speed controller, electric motor, and propeller. The reliability of the redundant system of propulsors is calculated from the union of the probabilities of each failure case (no failure, single failures, and double failures) that maintain full or degraded controllability of all the control axis. The co-axial quadrotor can keep control of all the control axis for any single failure.

Table 1 – All-axis controllability of the co-axial quadrotor with two propulsor failures.

Failed propulsors	1	2	3	4	5	6	7	8
1		X	X				X	
2				X				X
3				X		X		
4					X			
5						X		X
6							X	
7								X

X = Not controllable, i.e. propulsor system failure

Table 1 shows the controllability assessment for each combination of two propulsor failures. The combination of more than two failures is neglected, as they are too remote to affect the overall reliability and are not relevant for design consideration. The resulting reliability of the system of propulsors is integrated into the RBD shown in Figure 1 and expressed as follows:

$$R_{RSP}(t) = 16 \cdot R_{prop}(t)^6 \cdot \Big(1 - R_{prop}(t)\Big)^2 + 8 \cdot R_{prop}(t)^7 \cdot \Big(1 - R_{prop}(t)\Big) + R_{prop}(t)^8 \qquad (3)$$

Where, $R_{RSP}(t)$ is the reliability of the redundant system of propulsors, and $R_{prop}(t)$ is the reliability of a propulsor, which is the product of the speed controller, motor, and propeller reliabilities.

The failure rates shown in Figure 1 represent modern high-end transport category aircraft equipment, which is probably overly optimistic. Therefore, it is interesting to evaluate the impact of derating those failure rates. Table 2 shows the reliability of conventional and redundant design for failure rates rating from toy industry (derating 1000) to high-end transport category aircraft (no derating).

The results in Table 2 illustrate the potential of a fully redundant design, but implementation feasibility needs to be demonstrated and several challenges need to be addressed.

For example, the analysis assumes an ideal (no failure) battery management system, majority voting algorithms, electric wiring, and airframe. The actual implementation of these components could significantly impact the overall reliability as it can reveal single points of failure.

Table 2 – Emergency landing probability of failure

Failure rate derating	Conventional	Redundant
x1000 (toy industry)	6.42E-02	4.40E-03
x100	6.62E-03	4.65E-05
x10	6.64E-04	4.68E-07
x1 (transport category aircraft)	6.64E-05	4.68E-09

2.5.2. Conventional and redundant design sizing

Both conventional and redundant designs are sized with the approach described in Section 2.2. The conventional design uses a LiPo 12S-44.4V-25C-32000 mAh battery, while the redundant one uses a LiPo 12S-44.4V-25C-62000 mAh. The mass of redundant components and additional fittings, including harness, is also considered. The frame size remains the same for both designs, but the frame mass of the redundant design increases by +50% to accommodate redundant components and doubled battery capacity. It results in an additional mass of +18 kg (~82.5% increase) for the redundant design, as illustrated in Figure 2. It shows that the reliability and safety measures can significantly impact the multicopter mass.

2.5.3. Limitations

Due to the sizing tool limitations, the design used for the case study is a medium-sized midrange multicopter, while major safety-critical applications like medical transport and air-taxi would be significantly heavier.

As no sizing methodology integrating reliability is available, the presented work incorporates straightforward sizing and systematic redundancy measures without optimization. This approach can lead to solutions with insufficient performance in failure cases.

For simplicity, this study assumes that the components will fail without complex cascading effects. For example, it assumes that an electric motor failure will cancel its thrust, whereas it could instead overspeed or operate erratically and vibrate or overheat, leading to further failures.

The presented study is explorative and does not incorporate advanced safety considerations like hazardous and major failure conditions, generic failures, common cause failures, and particular risks. Possibly, this could worsen the complexity and weight impact on the design.

3. CONCLUSION

The literature review shows a lack of a comprehensive design for reliability methodology for UAV/AAV; in particular, there is no conceptual design methodology based on sizing with safety and reliability considerations.

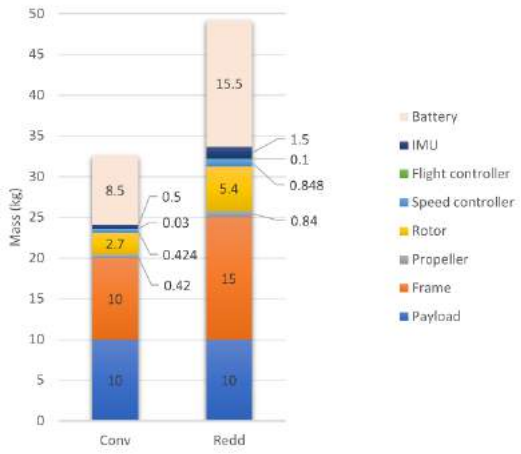


Figure 2. Mass comparison of conventional design (Conv) and redundant design (Redd)

Thus, this paper investigates the relevance of addressing this knowledge gap and the associated challenges to provide research directions. For the matter, it applies a straightforward approach combining redundancy in the sizing and couples controllability assessment with reliability calculation in a case study.

The case study compares a conventional with a redundant design. The simplicity of the approach avoids developing a new methodology and tool. However, it prevents optimizing the sizing and the redundancy jointly. Consequently, the redundant design is fault-tolerant and has the potential to meet the reliability requirements, but its mass is almost double that of the conventional design. Plus, the mass impact could further increase with the consideration of failure cases into sizing. This result confirms the relevance of developing a new conceptual design methodology based on a sizing including safety and reliability considerations. Towards such a methodology, future works should focus on bridging controllability analysis with sizing and reliability evaluation to optimize redundancy.

A future comprehensive design for reliability methodology should also consider more complex failure modes (e.g., motor overspeed, overheating, and vibration), hazardous and major failure conditions, and the effects of generic, common cause failures, and particular risks for completeness. Also, the feasibility of the redundancy techniques must be demonstrated with control and monitoring systems capable of detecting and containing failures before a critical destabilization of the vehicle.

Future works must answer all these challenges for an effective conceptual design of optimized UAV/AAV for safety-critical applications.

4. REFERENCES

[1] Q. Quan, *Introduction to multicopter design and control*. Springer, 2017.

[2] E. Lygouras, A. Gasteratos, K. Tarchanidis, and A. Mitropoulos, "ROLFER: A fully autonomous aerial rescue support system," *Microprocess. Microsyst.*, vol. 61, pp. 32–42, Sep. 2018, doi: 10.1016/j.micpro.2018.05.014.

[3] I. Nizar, Y. Illoussamen, H. El Ouarrak, E. Hossein Illoussamen, M. Grana (Graña), and M. Mestari, "Safe and optimal navigation for autonomous multi-rotor aerial vehicle in a dynamic known environment by a decomposition-coordination method," *Cogn. Syst. Res.*, vol. 63, pp. 42–54, Oct. 2020, doi: 10.1016/j.cogsys.2020.05.003.

[4] C. A. Ochoa and E. M. Atkins, "Fail-Safe Navigation for Autonomous Urban Multicopter Flight," presented at the AIAA Information Systems-AIAA Infotech @ Aerospace, Grapevine, Texas, Jan. 2017. doi: 10.2514/6.2017-0222.

[5] M. Gatti, "Complete Preliminary Design Methodology for Electric Multirotor," *J. Aerosp. Eng.*, vol. 30, no. 5, p. 04017046, Sep. 2017, doi: 10.1061/(ASCE)AS.1943-5525.0000752.

[6] T. Du, A. Schulz, B. Zhu, B. Bickel, and W. Matusik, "Computational multicopter design," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 1–10, Nov. 2016, doi: 10.1145/2980179.2982427.

[7] M. Biczyski, R. Sehab, J. F. Whidborne, G. Krebs, and P. Luk, "Multirotor Sizing Methodology with Flight Time Estimation," *J. Adv. Transp.*, p. 15.

[8] W. Ong, S. Srigrarom, and H. Hesse, "Design Methodology for Heavy-Lift Unmanned Aerial Vehicles with Coaxial Rotors," presented at the AIAA Scitech 2019 Forum, San Diego, California, Jan. 2019. doi: 10.2514/6.2019-2095.

[9] S. Delbecq, M. Budinger, A. Ochotorena, A. Reysset, and F. Defay, "Efficient sizing and optimization of multirotor drones based on scaling laws and similarity models," *Aerosp. Sci. Technol.*, vol. 102, p. 105873, Jul. 2020, doi: 10.1016/j.ast.2020.105873.

[10] G.-X. Du, Q. Quan, B. Yang, and K.-Y. Cai, "Controllability Analysis for Multirotor Helicopter Rotor Degradation and Failure," *J. Guid. Control Dyn.*, vol. 38, no. 5, pp. 978–985, May 2015, doi: 10.2514/1.G000731.

[11] D. Shi, B. Yang, and Q. Quan, "Reliability analysis of multicopter configurations based on controllability theory," in *2016 35th Chinese Control Conference (CCC)*, Chengdu, Jul. 2016, pp. 6740–6745. doi: 10.1109/ChiCC.2016.7554418.

[12] A. Chamseddine, I. Sadeghzadeh, Y. Zhang, D. Theilliol, and A. Khelassi, "Control Allocation for a Modified Quadrotor Helicopter Based on Reliability Analysis," presented at the Infotech@Aerospace 2012, Garden Grove, California, Jun. 2012. doi: 10.2514/6.2012-2511.

[13] Y. Zhang, V. S. Suresh, B. Jiang, and D. Theilliol, "Reconfigurable Control Allocation against Aircraft Control Effector Failures," in *2007 IEEE International Conference on Control Applications*, Singapore, Oct. 2007, pp. 1197–1202. doi: 10.1109/CCA.2007.4389398.

[14] X. Yu, L. Guo, Y. Zhang, and J. Jiang, *Autonomous Safety Control of Flight Vehicles*, 1st ed. First edition. | Boca Raton, FL : CRC Press, an imprint of Taylor & Francis Group, 2021.: CRC Press, 2021. doi: 10.1201/9781003144922.

[15] G.-X. Du, Q. Quan, and K.-Y. Cai, "Controllability Analysis and Degraded Control for a Class of Hexacopters Subject to Rotor Failures," *J. Intell. Robot. Syst.*, vol. 78, no. 1, pp. 143–157, Apr. 2015, doi: 10.1007/s10846-014-0103-0.

[16] H. Başak and E. Prempain, "Switched fault tolerant control for a quadrotor UAV," *IFAC-Pap.*, vol. 50, no. 1, pp. 10363–10368, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.1686.

[17] B. Wang, Y. Shen, and Y. Zhang, "Active fault-tolerant control for a quadrotor helicopter against actuator faults and model uncertainties," *Aerosp. Sci. Technol.*, vol. 99, p. 105745, Apr. 2020, doi: 10.1016/j.ast.2020.105745.

[18] P. Tang, D. Lin, D. Zheng, S. Fan, and J. Ye, "Observer based finite-time fault tolerant quadrotor attitude control with actuator faults," *Aerosp. Sci. Technol.*, vol. 104, p. 105968, Sep. 2020, doi: 10.1016/j.ast.2020.105968.

[19] J. Zha, X. Wu, J. Kroeger, N. Perez, and M. W. Mueller, "A collision-resilient aerial vehicle with icosahedron tensegrity structure," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, Oct. 2020, pp. 1407–1412. doi: 10.1109/IROS45743.2020.9341236.

[20] D. Floreano, S. Mintchev, and J. Shintake, "Foldable drones: from biology to technology," Portland, Oregon, United States, May 2017, p. 1016203. doi: 10.1117/12.2259931.

[21] *Flight evaluation of multicopter*. https://flyeval.com/.

[22] X. Dai, Q. Quan, J. Ren, and K.-Y. Cai, "An Analytical Design-Optimization Method for Electric Propulsion Systems of Multicopter UAVs With Desired Hovering Endurance," *IEEEASME Trans. Mechatron.*, vol. 24, no. 1, pp. 228–239, Feb. 2019, doi: 10.1109/TMECH.2019.2890901.

[23] X. Dai, Q. Quan, J. Ren, and K.-Y. Cai, "Efficiency Optimization and Component Selection for Propulsion Systems of Electric Multicopters," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 7800–7809, Oct. 2019, doi: 10.1109/TIE.2018.2885715.

[24] *MOC SC-VTOL- Proposed Means of Compliance with the Special Condition VTOL*, European union aviation safety agency. 2020. [Online]. Available: https://www.easa.europa.eu/document-library/product-certification-consultations/special-condition-vtol

[25] *Light Unmanned Aircraft Systems - SC Light-UAS 01*, European union aviation safety agency. 2020. [Online]. Available: https://www.easa.europa.eu/sites/default/files/dfu/special_condition_light_uas.pdf

[26] V. Lippiello, F. Ruggiero, and D. Serra, "Emergency landing for a quadrotor in case of a propeller failure: A backstepping approach," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, USA, Sep. 2014, pp. 4782–4788. doi: 10.1109/IROS.2014.6943242.

[27] A. Lanzon, A. Freddi, and S. Longhi, "Flight Control of a Quadrotor Vehicle Subsequent to a Rotor Failure," *J. Guid. Control Dyn.*, vol. 37, no. 2, pp. 580–591, Mar. 2014, doi: 10.2514/1.59869.

[28] Yu. V. Morozov, "Emergency Control of a Quadrocopter in Case of Failure of Two Symmetric Propellers," *Autom. Remote Control*, vol. 79, no. 3, pp. 463–478, Mar. 2018, doi: 10.1134/S0005117918030062.

[29] SAE International, "ARP 4761 - Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment," Aerospace Recommended Practice, Dec. 1996. doi: 10.4271/ARP4761.

ON SECURING CLOUD-HOSTED CYBER-PHYSICAL SYSTEMS USING TRUSTED EXECUTION ENVIRONMENTS

Amir Mohammad Naseri, Walter Lucia, Mohammad Mannan, Amr Youssef

Concordia Institute for Information Systems Engineering (CIISE)
Concordia University, Montreal, Canada

ABSTRACT

Recently, cloud control systems have gained increasing attention from the research community as a solution to implement networked cyber-physical systems (CPSs). Such an architecture can reduce deployment and maintenance costs albeit at the expense of additional security and privacy concerns. In this paper, first, we discuss state-of-the-art security solutions for cloud control systems and their limitations. Then, we propose a novel control architecture based on Trusted Execution Environments (TEE). We show that such an approach can potentially address major security and privacy issues for cloud-hosted control systems. Finally, we present an implementation setup based on Intel Software Guard Extensions (SGX), and validate its effectiveness on a testbed system.

Index Terms— Encrypted Control Systems, Trusted Execution Environments, Cloud/Edge-based CPS.

1. INTRODUCTION

With the development of cloud services, the implementation of industrial control systems into the cloud/edge has received increasing attention. The use of such services saves on the cost of setting up and maintaining industrial control systems (ICSs), as well as off-loading computationally expensive tasks. Moreover, when ICSs are geographically distributed, these cloud services are highly available and accessible from different locations [1]. When using cloud services in such applications, the main concern is the security and privacy of the cloud environment and communication channels between the physical plant and the cloud-hosted controller.

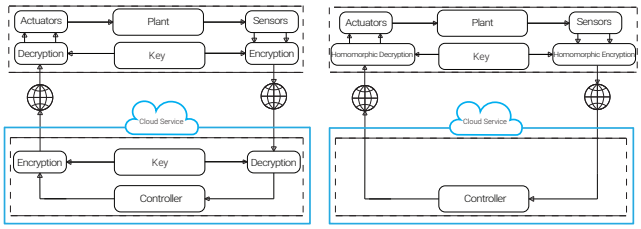
Different approaches have been proposed to enhance the security and privacy of networked CPSs where the controller is hosted in a cloud infrastructure. For example, Zhou et al. [2] propose the use of conventional cryptographic algorithms to secure plant-to-cloud communication. Kogiso and Fujita [3] propose the use of homomorphic encryption to ensure that the controller's operations can be performed without decrypting the received data, and hence addressing the confidentiality problem in the cloud (in addition to securing communication channels). Homomorphic encryption-based solutions have received increasing attention by the CPSs community; for full homomorphic and Pailier's homomorphic based solutions, see e.g., [4–7]. However, these homomorphic solutions suffer from unavoidable limitations related to the arithmetic operations allowed by the homomorphic schemes, ciphertext size explosion, and computation overhead. For solutions targeting only securing communication channels cannot protect controller logic and data against a malicious or compromised cloud provider. For data and execution security in the context of IoT and CPS applications, Shepherd et al. [8] survey and compare several existing secure and trusted computing environments such as Trusted Platform Module (TPM), Secure Elements (SE), Trusted Execution Environments (TEEs), and Encrypted Execution Environments (E3).

In this paper, we explore the use of encryption and trusted execution environments to secure plant-to-cloud communication channels and protect data and controller logic for cloud-hosted/edge-hosted CPS applications. To understand the performance implications of our approach, we also design and implement a simple prototype for a quadruple tank system [9], using Intel SGX as our TEE. Our results indicate that the introduced overhead is negligible, and highly scalable yet secure CPS applications can be designed for a cloud/edge-deployment scenario. We hope that our initial results may be useful to the CPS security community and encourage the design of more efficient and secure TEE-based solutions compared to current schemes that rely mostly on conventional cryptographic mechanisms and homomorphic schemes.

2. SYSTEM SETUP AND THREAT MODEL

A typical cloud-based, networked control system consists of the following main components: the plant, the controller, the cloud, and the communication channels. The *plant* is the physical entity that we want to control. It is usually equipped with a set of *actuators* and *sensors*. The *controller* collects the sensor measurements and computes, according to a predefined control logic, the control commands sent to the actuators. In a cloud-based networked setup, the controller and the plant are spatially distributed, and the controller logic is implemented in a *cloud service* provider. The communication channels are used for a real-time and bi-directional ex-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE



(a) Encrypted communications. (b) Homomorphic encryption.

Fig. 1: Existing security solutions for cloud-based CPSs.

change of data (e.g., sensor measurements and control inputs) between the plant and the controller.

Threat Model. We consider the following attacks that can affect the privacy/security of the cloud-based CPS controllers.

Attacks against the communication channels - By adopting the conventional Dolev-Yao threat model [10], a malicious entity with access to the public communication channels is assumed to be able to eavesdrop on the transmitted data and/or modify their content. Therefore, potentially, the confidentiality and the integrity of the control system could be compromised. Indeed, such attackers can exploit the eavesdropped data to gain further information about the controlled system's behaviour and use their disruptive capabilities to launch sophisticated undetectable attacks such as replay, covert, zero-dynamics attacks [11, 12].

Attacks against the cloud service - If the cloud operator is malicious, or if the service is vulnerable, then an unauthorized entity (e.g., malware authors) might be able to gain access to the data transmitted between the plant and the controller, even if encrypted and authenticated communications are used. Indeed, such attackers could read the encryption key (key-management problem), intercept the transmitted data after decryption, and change the control logic (with the consequence of jeopardizing the whole control loop).

3. EXISTING SOLUTIONS

Different schemes have been proposed to secure networked control systems. A common solution is to use encrypted authenticated communications between the plant and the controller [13]; see Fig. 1a. Such a solution, at the cost of increased computational power to perform encryption/decryption operations at both the plant and controller's sides of the CPS, can mitigate the privacy and security issues related to cyber-attacks against the communication infrastructure. On the other hand, it does not address the security and privacy risks associated with the controller's deployment inside the cloud.

The use of homomorphic encryption has also been proposed to secure CPS solutions [3, 14]; see Fig 1b. A distinctive capability of such a solution is that it allows the controller to implement the control logic (in terms of additions and multiplications operations) directly on the received encrypted sensor measurements. Consequently, such an approach has the advantage of securing the communications while solving the privacy issues associated with the cloud infrastructure. However, common drawbacks of homomorphic encryption include: the mathematical operations performed on the encrypted data are typically limited and computationally expensive; and the plaintext to ciphertext bit expansion factor is usually very high. Consequently, homomorphic-based solutions might not be practical for securing industrial control systems with fast sampling rates or narrow bandwidth.

There are three different types of homomorphic encryption schemes, namely partially homomorphic encryption (PHE), somewhat homomorphic encryption (SHE), and fully homomorphic encryption (FHE). Each subclass is characterized by the set and number of encrypted operations allowed. Therefore, according to the limitations imposed by the used scheme, it might be challenging to recast any existing control algorithm into its encrypted counterpart. For example, FHE allows an unlimited number of encrypted addition and multiplication operations. Therefore it is particularly appealing to implement sophisticated control solutions such as dynamic feedback control or model predictive control. However, such freedom comes with a computational expensive bootstrapping process that makes FHE impractical to most control systems. Kim et al. [4] propose FHE to implement a dynamic output feedback controller using multiple controllers to avoid the bootstrapping delay. However, another inherent issue with FHE is that the ciphertext expansion might be up to $10000:1$ for an acceptable level of security of 100 bits [15]. Pailier's homomorphic encryption (PHE, supporting only encrypted additions) has also been proposed to implement a variety of controllers [5, 6]. However, due to memory issues related to the state of the dynamic encrypted controller (i.e., the number of bits required for its representation grows linearly with the number of iterations), the solution is usually limited to the use of resetting dynamics control laws. On the other hand, if a proportional controller is used, then, the control gain must satisfy some restrictive conditions imposed by the number of available bits [7].

Overall, existing solutions pose several limitations in terms of security/privacy/deployability to networked control systems. Moreover, no solutions have been proposed to protect CPSs against a malicious cloud operator, or malware that might be able to compromise the integrity of the control algorithm running on the cloud server.

4. OUR PROPOSAL

The objectives of our proposal are: secure the cloud-based CPSs against all the cyber-threat discussed in Section 2, and reduce the impact on the design and implementation of existing control strategies. The proposed secure control architecture has two essential components (see Fig. 2): an authenticated encryption scheme for securing the communication channels, and a TEE where the control logic is executed and the secret cryptographic keys, used by the authenticated encryption scheme, are stored.

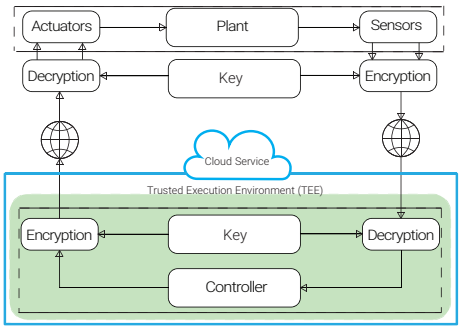


Fig. 2: Proposed TEE-based solution.

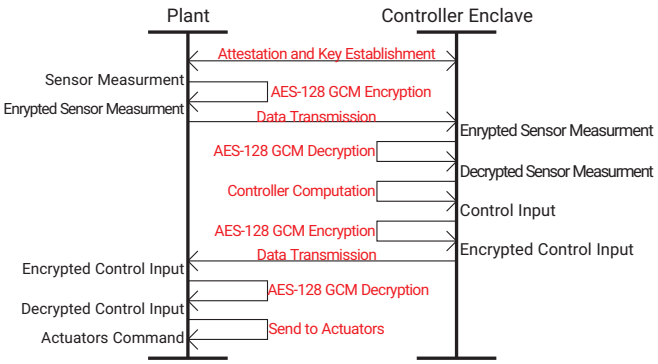


Fig. 3: Data flow in the proposed solution.

First, we resort to authenticated encryption schemes (cf. [13]) to ensure the integrity and the confidentiality of the control signal and sensor measurements exchanged between the plant and the controller. The used encryption scheme must be characterized by an inherent latency much smaller than the control-loop sampling time. The latter requirement is essential to ensure that the encryption scheme does not affect the control-loop system's stability. Second, a trusted execution environment (TEE) is used to protect the controller's operations in the cloud service. Generally speaking, a TEE refers to a hardware-based solution capable of ensuring that no malicious cloud entities (e.g., malware or a malicious cloud operator) could interfere with the execution of the control algorithm or with the memory associated with it. Moreover, if encryption/decryption operations are executed inside the TEE, where the keys are also protected by the TEE, then a malicious cloud administrator also cannot access the keys. TEE may also provide some other advantages such as measuring the integrity of the launched processes, measuring the origin of the TEE and current state of the TEE (attestability), and recovering the state of the TEE to a known good state after any corruption (recoverability). The presence of a TEE on the plant side is not required for our threat model. However, it is desirable in a scenario where the local computing platform (e.g., SCADA system) could be subject to cyber-attacks. Several solutions have been proposed in the literature (not in CPS) using different TEE implementations, e.g., Intel SGX [16], ARM TrustZone [17], AMD SEV [18], Hardware Security Module (HSM) [19], and secure co-processors [20]. Although all these solutions provide strong security mechanisms, not all can be used in our design (e.g., HSMs do not support remote attestation as opposed to Intel SGX).

5. IMPLEMENTATION

We use Intel SGX as TEE for its capability of providing a cryptographic attestation to ensure the integrity of the execution of the controller algorithm, even in the presence of a malicious cloud admin or a compromised cloud operating system (e.g., by a malware). To keep code and data secure, SGX provides an isolated execution environment, and encrypted memory. This secure container is called an "enclave" and everything else outside the enclave is assumed to be insecure. Two main functions are available to interact with the enclave, namely Enclave Call (E-CALL) and Out Call (O-CALL). E-CALL is used to call, from outside the enclave, a function implemented inside the enclave. On the other hand, O-CALL is used to call, from inside the enclave, a function implemented outside the enclave. For the implementation of the authenticated encryption, AES-128 Galois/Counter Mode(GCM) is used. This algorithm is a good candidate for CPSs because of its high throughput and low latency [21, 22]. First, we need to create an enclave and allocate memory for the Enclave Page Cache (EPC). The process starts with the attestation of both the enclave (validity of the CPU's SGX support) and the code (validity of the binary executed within the enclave as the controller logic). During the attestation, entities also establish a secure session key. After these initialization operations, data transmission will be started between the participating entities, encrypted under the session key. The data flow for a single control loop is shown in Fig. 3. In particular, the sensor measurements are encrypted on the plant side. Then, these encrypted sensor measurements are sent to the cloud over the communication channel. The authenticity of the received measurement is checked inside the enclave, where then the controller logic is also applied to the decrypted measurements. The evaluated controller output is then encrypted (inside the enclave) before it is sent to the actuator through the communication channel. Finally, the encrypted control input is decrypted by the actuator and applied to the plant.

6. SECURITY AND PERFORMANCE EVALUATION

We now discuss the security properties of the proposed solution. (i) *Confidentiality:* Data sent through the communication channels are encrypted with AES-128 GCM. Therefore, network eavesdroppers are unable to decrypt the transmitted control signals and sensor measurements. Moreover, control operations and encryption/decryption operations are performed within the enclave, avoiding the possibility that a malware or cloud administrator could intercept the plaintext signals or acquire the keys. (ii) *Integrity:* By exploiting the

message authentication code (MAC) tag in AES-128 GCM, it is possible to verify the integrity of the transmitted data (i.e., detect if an attacker has manipulated the transmitted data). Another aspect of integrity is to make sure that the controller logic is not manipulated by the cloud provider before the code is executed within the enclave. For this purpose, an attestation operation is performed to make sure that the code executed in the enclave is exactly that is sent to the cloud service by the system admin. To improve code obfuscation (i.e., hiding the control logic from the cloud operator), the proposed solution in [23] can be used. Note that the controller's runtime state remains always protected by SGX's memory encryption. Moreover, since the controller is executed inside SGX, the integrity of the control algorithm is also ensured. (iii) *Authentication:* The remote attestation feature of Intel SGX is used on the plant side to establish a secure and authenticated communication channel with the enclave in the cloud and ensure that the remote enclave is trusted. The MAC tags also is used by both entities (plant and controller) to make sure that the received messages are obtained by a trusted entity. (iv) *Freshness:* The uniqueness of the AES-128 GCM IV is used to guarantee freshness of each message. Defending against side-channel attacks against Intel SGX [24] is outside the scope of this paper. In the case of necessity of storing data by the controller (depend on the controller logic), to mitigate rollback attacks on the sealed data, Monotonic Counter (MC) of Intel SGX can be used to guarantee that the sealed data is the latest copy.

6.1. Performance Evaluation

System setup. As a testbed, we use the Quadruple Tank Process (QTP) system from Johansson [9], which is often used as a benchmark for control systems applications. The system consists of four water tanks where $h_i, i \in 1, 2, 3, 4$ represents the level of water in each tank and also represents the states x of the system, i.e., $x = [h_1,\, h_2,\, h_3,\, h_4]^2 \in \mathbb{R}^4$. There are two sensors that measure the level of water inside tanks 1 and 2, i.e., the output measurement vector is $y = [0.5h_1,\, 0.5h_2]^T \in \mathbb{R}^2$. Moreover, the system is equipped with two pumps and the applied voltage v_1, v_2 are the inputs u of the system, i.e., $u = [v_1,\, v_2]^T$. We have linearized the system model around the equilibrium pair ($x_{eq} = [12.4,\, 12.7\,, 1.8,\, 1.4]^T$ $u_{eq} = [3,\, 3]^T$) and discretized it using a sampling time $T_s = 0.1\,\text{sec}$. The linearized model $x(k+1) = Ax(k) + Bu(k)$, $y(k) = Cx(k)$ and its matrices A, B, C can be easily obtained following [9]. The plant is regulated by means of dynamic output feedback controller consisting of a Luenberger Observer and an optimal Linear Quadratic (LQ) controller. The state-estimator operations are described by the discrete-time system $\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + L(y(k) - C\hat{x}(k))$ where $\hat{x}(k)$ is the estimation of the state $x(k)$ and the correction gain is given by $L = \begin{bmatrix} 0.78 & 0 & 0.32 & 0 \\ 0 & 0.78 & 0 & 0.32 \end{bmatrix}^T$. The LQ controller logic is computed as $u = K(x - x_{eq}) + u_{eq}$ where the stabilizing gain is given by $K = \begin{bmatrix} 27.547 & -0.054 & 0.468 & 0.086 \\ 0.023 & 28.441 & 0.143 & 0.507 \end{bmatrix}$. The dynamic output feedback controller operations have been implemented by utilizing an Intel SGX running on an Intel Core i7-6700 CPU, 3.40GHz, with 4 cores and 8 threads and 16 GB of RAM, using 64-bit Windows 7.

Measurements. We have conducted a series of measurements to evaluate the computation times required by different components of the proposed solution (see the data flow in Fig. 3). The reported CPU measurements have been obtained using the approach proposed in [25, Fig. 1], i.e., an O-CALL function is used as a stopwatch. As a result, the time measurements in Table 1 include an extra time representing the CPU time required to return to the enclave from an O-CALL and exit from it. We denote this time by Δt. To mitigate the presence of Δt in the measurements, we repeated each operation inside the enclave 1000 times and then calculate the average. Δt is also measured separately. The numerical results show that the two dominant factors are Δt and the control algorithm CPU time. Indeed, the average total CPU time required by both the secure and insecure implementations are around $905\mu s$ and $479\mu s$, respectively. The obtained results confirm that the computational overhead introduced by the use of Intel SGX does not affect the feasibility of the control strategy. Moreover, given that the introduced overhead is in the milliseconds' range, the proposed SGX-based secure architecture is believed to be affordable for a large class of cloud-based control systems applications.

Operation	Time (μs)
Enclave creation	8368.4
Dynamic output feedback controller	466.7
AES-128 GCM encryption	1.8
AES-128 GCM decryption	1.4
Δt	435.4

Table 1: Average time for different operations of the SGX-based solution

7. CONCLUSION

We proposed a solution to secure cloud-hosted/edge-hosted CPSs. In particular by resorting to authenticated encryption and a trusted execution environment, we showed that the proposed networked control scheme is secure again different attacks against its security and privacy. We verified the effectiveness of such a scheme by means of numerical simulations obtained considering Intel SGX, where we performed different benchmarks to evaluate the computational burden associated to the trusted control scheme implementation. The obtained results show good promise in terms of real-time performance and simplicity of implementation in CPSs applications. The proposed solution can also be implemented in a non cloud setting to help mitigating supply chain breaches.

8. REFERENCES

[1] M. S. Mahmoud and Y. Xia, *Networked control systems: cloud control and secure control*, Butterworth-Heinemann, 2019.

[2] L. Zhou, V. Varadharajan, and M. Hitchens, "Achieving secure role-based access control on encrypted data in cloud storage," *IEEE trans. on information forensics and security*, vol. 8, no. 12, pp. 1947–1960, 2013.

[3] K. Kogiso and T. Fujita, "Cyber-security enhancement of networked control systems using homomorphic encryption," in *IEEE Conf. on Decision and Control (CDC)*. IEEE, 2015, pp. 6836–6843.

[4] J. Kim, C. Lee, H. Shim, J. H. Cheon, A. Kim, M. Kim, and Y. Song, "Encrypting controller using fully homomorphic encryption for security of cyber-physical systems," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 175–180, 2016.

[5] J. Tran, M. Farokhi, F.and Cantoni, and I. Shames, "Implementing homomorphic encryption based secure feedback control," *Control Engineering Practice*, vol. 97, pp. 104350, 2020.

[6] C. Murguia, F. Farokhi, and I. Shames, "Secure and private implementation of dynamic controllers using semi-homomorphic encryption," *IEEE Trans. on Automatic Control*, vol. 65, no. 9, pp. 3950–3957, 2020.

[7] Y. Lin, F. Farokhi, I. Shames, and D. Nešić, "Secure control of nonlinear systems using semi-homomorphic encryption," in *IEEE Conf. on Decision and Control*, 2018, pp. 5002–5007.

[8] C. Shepherd, G. Arfaoui, I. Gurulian, R. P. Lee, K. Markantonakis, D. Akram, R. N.and Sauveron, and E. Conchon, "Secure and trusted execution: Past, present, and future-a critical review in the context of the internet of things and cyber-physical systems," in *IEEE Trustcom/BigDataSE/ISPA*, 2016, pp. 168–177.

[9] K. H. Johansson, "The quadruple-tank process: A multivariable laboratory process with an adjustable zero," *IEEE Trans. on control systems Tech*, vol. 8, no. 3, pp. 456–465, 2000.

[10] D. Dolev and A. Yao, "On the security of public key protocols," *IEEE Trans. on information theory*, vol. 29, no. 2, pp. 198–208, 1983.

[11] S.M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, "A systems and control perspective of CPS security," *Annual Reviews in Control*, 2019.

[12] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[13] S.C. Patel, G. D. Bhatt, and J. H. Graham, "Improving the cyber security of SCADA communication networks," *Communications of the ACM*, vol. 52, no. 7, pp. 139–142, 2009.

[14] F. Farokhi, I. Shames, and N. Batterham, "Secure and private cloud-based control using semi-homomorphic encryption," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 163–168, 2016.

[15] I. Chillotti, N. Gama, M. Georgieva, and M. Izabachène, "Tfhe: fast fully homomorphic encryption over the torus," *Journal of Cryptology*, vol. 33, no. 1, pp. 34–91, 2020.

[16] V. Costan and S Devadas, "Intel SGX explained.," *IACR Cryptol. ePrint Arch.*, vol. 2016, no. 86, pp. 1–118, 2016.

[17] AMBA Infrastructure, "Technical overview," 2004.

[18] D. Kaplan, J. Powell, and T. Woller, "Amd memory encryption," *White paper*, 2016.

[19] J. Varia, S. Mathew, and et. al, "Overview of amazon web services," *Amazon Web Services*, vol. 105, 2014.

[20] S. Bajaj and R. Sion, "Trusteddb: A trusted hardware-based database with privacy and data confidentiality," *IEEE Trans. on Knowledge and Data Engineering*, vol. 26, no. 3, pp. 752–765, 2013.

[21] Sandhya Koteshwara, Amitabh Das, and Keshab K Parhi, "FPGA implementation and comparison of AES-GCM and Deoxys authenticated encryption schemes," in *IEEE Int. symposium on circuits and systems*, 2017, pp. 1–4.

[22] V. Arun, K. Vanisree, and D. Reddy, "Implementation of AES-GCM encryption algorithm for high performance and low power architecture using FPGA," 2015.

[23] E Bauman, H Wang, M Zhang, and Z Lin, "Sgxelide: enabling enclave code secrecy via self-modification," in *Proceedings of Int. Symposium on Code Generation and Optimization*, 2018, pp. 75–86.

[24] F. Brasser, U. Müller, A. Dmitrienko, S. Kostiainen, K.and Capkun, and A.-R. Sadeghi, "Software grand exposure: SGX cache attacks are practical," in *USENIX Workshop on Offensive Tech.*, 2017.

[25] A. T. Gjerdrum, R. Pettersen, H. D. Johansen, and D. Johansen, "Performance principles for trusted computing with intel SGX," in *Int. Conf. on Cloud Computing and Services Science*. Springer, 2017, pp. 1–18.

A STRESS TESTING FRAMEWORK FOR AUTONOMOUS SYSTEM VERIFICATION AND VALIDATION (V&V)

Gregory Falco *

Johns Hopkins University
Institute for Assured Autonomy
Baltimore, MD

Leilani H. Gilpin

Massachusetts Institute of Technology
CSAIL
Cambridge, MA

ABSTRACT

Autonomous cyber-physical systems are prone to error and failure. Verification and validation (V&V) is necessary for their safe, secure and resilient operations. Methods to detect faults in aerospace engineering (fault trees) and later adapted for security (attack trees) could capture a wide array of critical risks and we argue how stress testing could be a pragmatic approach to evaluating the assurance of autonomous cyber-physical systems.

***Index Terms*—** Stress testing, Autonomous Systems, Formal Methods, Cyber-physical systems, Robust AI, XAI, Assured autonomy, Verification and Validation, V&V

1. INTRODUCTION

Cyber-physical autonomous systems are prone to failures and are not currently tested properly. Verification and validation (V&V) testing must fully capture both physical safety and digital security risks, which are compounded by the inherent complexity of autonomous systems. Current V&V testing and proving properties can harden these systems, but they are inadequate–it is impossible to "formally" test all failure modes. The key idea is that these failures are not isolated. Instead of building provable properties, our research is a complementary approach: we propose work on *AI stress testing*.

Stress testing is crucial for autonomous cyber-physical systems in *open environments*. Image recognition systems have been shown to be brittle and biased [1], and this is illuminated as a threat to humanity in the domain of self driving cars [2]. These mistakes and errors need to become test cases, similar to the types of stress testing that is done in consumer vehicles, aerospace systems, and commercial aircrafts. We discuss the merits of stress testing via a risk-based approach to build trust and security in autonomous, cyber-physical systems. While a stress test should be customized to the system of interest, we propose a consistent approach to evaluating and interpreting the results of stress tests to successfully compare V&V tests across autonomous agents. Our stress test evaluation framework is based on methods that have been in use for decades in safety science. We provide an example for how our stress testing framework could be employed for the autonomous agents that comprise NASA's future lunar habitat - the Artemis Base Camp.

*This research was supported in part by the Fulbright Commission and the National Science Foundation.

2. PRIOR WORK ON V&V FOR AUTONOMY

Safety-critical systems need appropriate testing protocols. Human operators of machinery or personal vehicles are subject to driving tests, safety protocols, and certifications. Autonomous operators should be subject to the same types of testing.

But what do we seek to understand from these tests? There has been work on documenting failures, but there is an increasing need to categorize and prioritize autonomous system needs and challenges [3]. The AI incident database [4] was released as a means to avoid "repeated AI failures [by] making past failures known." We are inspired by the work of the AI incident database to distill past failures into an accessible testing framework. There have been many V&V mechanisms proposed for autonomous agents[5]. Below is a small sampling of some predominant tests for autonomous agents, each of which have notable draw-backs.

Formal methods is among the most used V&V testing techniques that has been employed for safety-critical systems [6, 7]. However, there are certain characteristics of autonomous agents that are not conducive to formal methods. For example, autonomous agents generally lack "unambiguous" requirements and specifications, they operate in semi-known environments that may change at a moment's notice, and they may hand off control to a human operator at some point in the mission thereby introducing further uncertainty into the operating equation [8]. Additionally, there is often incomplete information about what went into the training of the agent and its subsequent learned behavior. The agent may have learned "unsafe" behavior, unknown to operators [8].

There are also challenges using formal methods to eval-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

uate the security of an autonomous agent. Many have tried to remedy formal methods for autonomous applications [9, 10, 11], including work that is quite similar to our contribution: using some sort of fault tree to derive verification properties [12, 13]. But, formal methods has struggled to gain traction in security testing communities, given the ever-expanding state space and unpredictability of creative attackers. For example, formal methods will not be able to detect a potential issue associated with previously unseen vulnerabilities or exploits [14]. This is the very reason why many security researchers still employ attack trees rather than formal methods to evaluate security holes in complex systems. Ultimately, the challenge with formal methods is that they are generally reliant on specifications, static analysis, well-known outcomes and determinism to develop a strong model - whereas autonomous agents change at run-time given that they are constantly learning and making decisions in undefined environments.

Differential testing is generally engaged to make sure that different versions of software that may have been updated produce a consistent output [15]. It has been used for both cyber-physical systems and information technology systems alike. A challenge engaging this approach for autonomous agents is that it only intends to capture changes in operation between different versions - not identify net new risks.

Simulation testing is commonly employed in reinforcement learning, where the agent training process involves sequential Markov decision problems which act as essentially a series of stress tests. Algorithms that can be engaged for this simulation include a Monte Carlo tree search or deep reinforcement learning[5]. Usually, these "tests" occur in a realistic, but closed-world simulation. The problems arise with this approach when these agents transfer to real, open world environments given their dependence on some pre-existing domain knowledge which can be poorly defined in unknown environments.

3. FAILURE TYPES AND THEIR STRESSORS

There are three failure axes for cyber-physical systems. The system can fail due to an internal fault (in Section 3.1), or an error that can be pinpointed to a part or connection inside the system. Another failure mode is due to an unexpected external factor (in Section 3.2); an attack or one-off incident from external factors, such as weather. Finally, a less considered, but equally important failure mode in the context of testing is that of ethics (in Section 3.3). Autonomous agent ethics has been robustly discussed for autonomous agents [16], but less so in the context of testing.

For each axis, we propose a series of stressors that induce the associated failures. The stressors should be individually tested for each autonomous agent. The specific tests employed for the stressors should vary depending on the type of agent being stress tested; however the tests should be evaluated in a consistent manner so that systems engineers can compare and prioritize failures.

Importantly, the questions aim to distinguish between failures that matter in the context of autonomous agent resilience and others that do not. Autonomous agents are inherently complicated and will therefore be prone to failures - but not all will be consequential. Stress tests should elucidate this distinction between failure severity. Resilience is used as the baseline requirement for distinguishing what failures matter because it indicates what failures an agent could tolerate while still achieving its mission. The questions are explicitly described further in the Stress Testing Evaluation Framework.

3.1. Internal Fault

Internal faults can be caused by stresses due to a failed component or a failed connection between parts. One type of local failure is a mechanical failure such as a sensor failure. This occurs when a mechanical component is obfuscated, misaligned, misinterpreted or malfunctions altogether. An obfuscation example is LiDAR sensors that cannot detect objects in the rain or snow [17]. Since sensor data is commonly noisy, it can be easily misinterpreted, which happens in wireless networks, vehicles, and other smart systems. And finally, sensors, like all subsystems can malfunction or crash. The main commonality between these failures are that they are *local* to the sensor subsystem.

Software bugs are another stress that can result in an internal fault, which can be local or between components. An example is the NaN error in the autonomous racecar[1], or the hallucinating behavior of deep network networks[1], which can be monitored with commonsense data and rules [18]. Other communication failures can be due to network latency, incorrect assumptions, or other external factors, which we cover in the next section.

3.2. External Forces

External forces on an agent could induce a variety of failures. One such external force is that of a cyberattack. Autonomous cyber-physical systems have a great deal of surface area that could be subject to attack. Attackers may be particularly attracted to autonomous agents given the grandeur and physical impact of their potential failure. Attackers can target anything from the training data set to the control system itself. Cyber-physical autonomous agents are finely tuned where even a slight timing attack could throw off the real-time operating systems inherent to these agents. A timing attack to an autonomous robotic arm operating in a chemical plant could cause an explosion should chemical compounds be mixed at the incorrect frequency. While not a fully autonomous agent,

[1] Autonomous racecar slams into a wall: https://www.thedrive.com/news/37342/autonomous-race-car-starts-test-lap-immediately-slams-into-wall

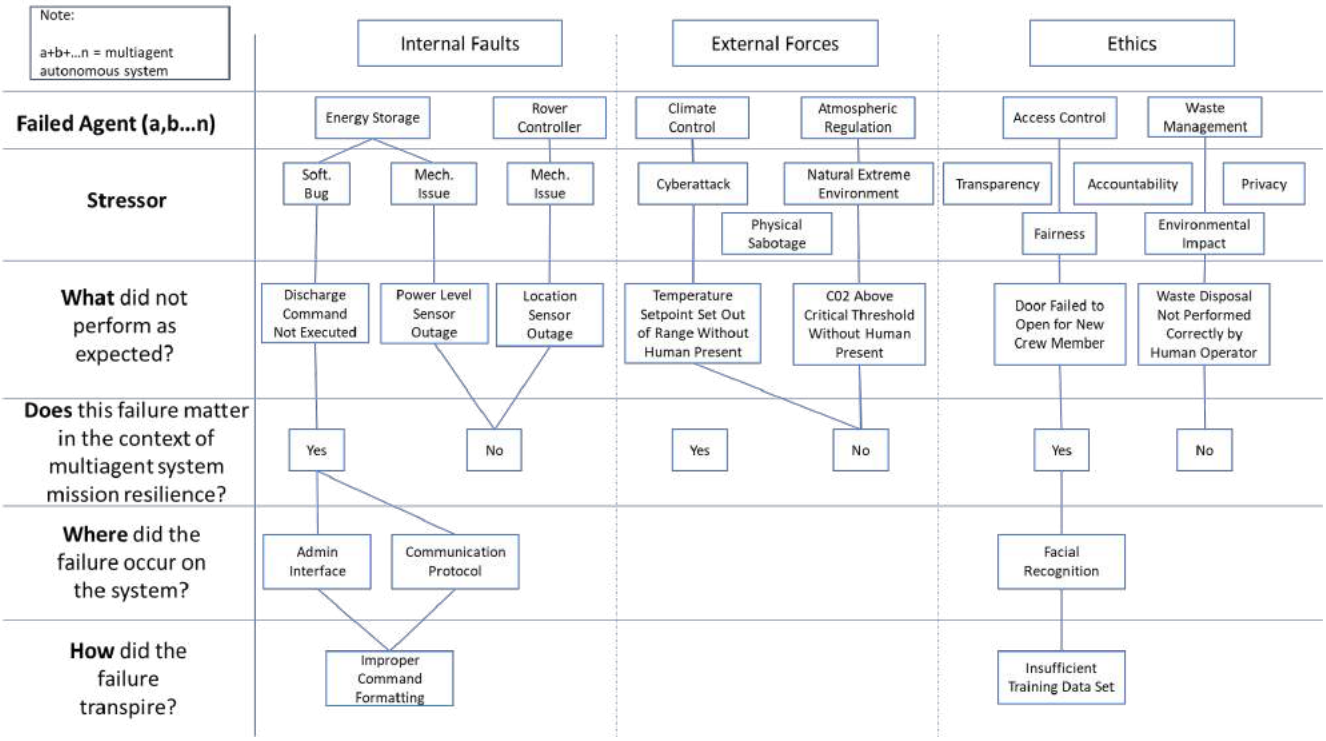


Fig. 1. Our stress test evaluation framework.

a similar cyber incident occurred at a German steel mill in 2014 causing significant damage to the plant [19].

The physical nature of cyber-physical autonomous agents also poses the risk of physical sabotage. Drones are increasingly autonomous and being employed for important tasks such as in military surveillance and reconnaissance missions. There have been incidents where semi-autonomous drones have been shot down such as the Global Hawk Spy Drone by Iran in 2019 [20].

A less considered, but equally devastating external force is natural extreme environmental conditions, such as weather. Many autonomous agents are designed to operate in extreme environments so that humans do not have to be present. An example of such an autonomous robotic agent is one used for deep sea arctic exploration [21]. Autonomous agents have challenges with far less extreme environments - such as in rain, wind or fog, which have been demonstrated to induce mission failure in autonomous vehicles [22]. Such extreme natural conditions' impact can become compounded in autonomous agents - inciting system failure.

3.3. Ethics

An ontology of ethics stressors have been previously enumerated to include: transparency, accountability, privacy, fairness [16]. Each has the capacity to cause a failure that inhibits a system's mission resilience. An additional ethics stressor that has not been as discussed is environmental impact. Specifically, this could include how a system's performance may damage its surrounding environment while achieving its mission. For example, an autonomous robot whose mission is to retrieve a series of artifacts from a delicate environment such as an archaeological excavation may succeed in retrieving the artifact at the expense of the surrounding environment that housed the artifact - thereby inhibiting its ability to return to retrieve further specimens. This presents an ethical failure of the autonomous agent.

4. STRESS TESTING EVALUATION FRAMEWORK

We propose a hierarchical tree structure that serves to aid systems engineers to evaluate each agent's stressors across an autonomous system. This hierarchy employs the framework established for fault tree analysis (originally developed for the aerospace community in the 1960s) [23], which has been used extensively in the field of safety science and then later adapted by the security community in the form of attack trees [24]. Tree structures have been used to enumerate risk for automotive reliability and safety studies[25]. Generally these tree structures do not have significant structural requirements beyond enumerating subsequent detail as one descends the tree on how a component failed or is attacked. However, by furnishing each tree "branch" level with a series of questions about the failure, the systems engineer can more easily

compare and prioritize the failures for each agent. Establishing further structure for the tree hierarchy has been previously demonstrated [26].

5. EXAMPLE SCENARIO

To demonstrate how the stress testing evaluation framework could be employed, a sample is illustrated in Figure 1 concerning NASA's future autonomous lunar habitat. The scenario illustrates an autonomous agent that has been stress tested for each stressor for each autonomous agent described in Section 3. The framework would have been completed by a systems engineer after the stress test for each agent. A systems engineer could use any level of the tree hierarchy (question) as their prioritization filter; however, the failures that affect mission resilience should be addressed first.

The lunar habitat will be composed of a series of autonomous control system agents that will be required to work together with other agents and humans. In some cases agents will be acting with humans present and co-operated, while at other times the agents will be acting without the physical presence of humans. In all cases, the agents will be working towards the mission of establishing a sustained habitable environment that enables scientific exploration on the lunar surface. Agents that compose the autonomous lunar habitat may include, but is not limited to: resource (water, energy, materials, etc.) harvesting, resource (water, energy, materials, etc.) management (storage, allocation, discharge, etc.), vehicle control, vehicle maintenance, climate control, atmospheric regulation, access control, and waste management. The success of the mission will be reliant on the accomplishment of each agent's operations as well as their interactions. For example, a vehicular control system will be dependent on the resource management system given a lunar rover will require proper energy storage, allocation and distribution. The lunar habitat will exist in an inherently extreme environment with considerable failure risks from external forces. Given the complexity of the autonomous agents, there are also many internal faults that can possibly occur. The necessary agent-human and agent-environment interaction also poses the opportunity for ethical failures. Each stressor must be evaluated in the context of the operating parameters of the autonomous agent at any given time. Evaluating NASA's future lunar habitat is an especially interesting and critical case for stress testing given the lack of physical access to devices, extreme costs associated with repairs and the delicate nature of the overall mission. One autonomous agent's failure could ostensibly cause the lunar habitat to fail.

6. DISCUSSION AND FUTURE WORK

Although there has been previous work on documenting and classifying failure cases, there has been little work on what information is sought when a system fails. In this paper we have shown a proof-of-concept stress testing framework for cyber-physical autonomous agents. This is especially important for *assured autonomy* and building trust in our autonomous counterparts.

As autonomous agents take control of operation that was previously entrusted to humans, it is necessary to test these mechanisms in the same way that human operators are tested. With the increasing number of connections, parts, and complexity of these systems, the state space has evolved making it challenging to fully address using formal methods. Unlike other V&V frameworks, our approach offers a means for flagging issues without extensive data or quantitative analysis (which may be unavailable). The stress testing framework can customized to prioritize stressors and their associated failures to help ensure the autonomous agent's assurance. While some existing V&V methods are useful for static systems, it is time for the community to expand how autonomous agents are evaluated and stress testing will be a critical aspect of this. Now, it is imperative that we start testing and refining stress test evaluation frameworks such as the one proposed to help build trust in autonomous agents.

7. CONCLUSION

In this paper, we have revisited themes from classical fault diagnostics to chart a path forward for stress testing autonomous cyber-physical systems. Our stress testing framework enables end users to determine what they should be testing for (given each system is unique), while leaving it up to the systems engineers to devise sufficient tests for their systems. We do not believe that the stress testing framework proposed is comprehensive and we encourage the community to build on this to propose new questions critical to mission resilience and system assurance. Fundamentally, there is merit to strategically breaking the autonomous agent and methodically questioning and documenting what went wrong.

8. REFERENCES

[1] Anh Nguyen, Jason Yosinski, and Jeff Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, 2015, pp. 427–436, IEEE.

[2] Alexey Kurakin, Ian Goodfellow, and Samy Bengio, "Adversarial examples in the physical world," *arXiv preprint arXiv:1607.02533*, 2016.

[3] Gregory Falco, "Autonomy's hierarchy of needs: Smart city ecosystems for autonomous space habitats," in *2021 55th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2021.

[4] Sean McGregor, "Preventing repeated real world ai failures by cataloging incidents: The ai incident database," 2020.

[5] Anthony Corso, Robert J Moss, Mark Koren, Ritchie Lee, and Mykel J Kochenderfer, "A survey of algorithms for black-box safety validation," *arXiv preprint arXiv:2005.02979*, 2020.

[6] Shaoying Liu, Victoria Stavridou, and Bruno Dutertre, "The practice of formal methods in safety-critical systems," *Journal of Systems and Software*, vol. 28, no. 1, pp. 77–87, 1995.

[7] Yingxu Wang, Ming Hou, Konstantinos N Plataniotis, Sam Kwong, Henry Leung, Edward Tunstel, Imre J Rudas, and Ljiljana Trajkovic, "Towards a theoretical framework of autonomous systems underpinned by intelligence and systems sciences," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 1, pp. 52–63, 2020.

[8] Ufuk Topcu, Nadya Bliss, Nancy Cooke, Missy Cummings, Ashley Llorens, Howard Shrobe, and Lenore Zuck, "Assured autonomy: Path toward living with autonomous systems we can trust," *arXiv preprint arXiv:2010.14443*, 2020.

[9] Kristin Yvonne Rozier, "Specification: The biggest bottleneck in formal methods and autonomy," in *Working Conference on Verified Software: Theories, Tools, and Experiments*. Springer, 2016, pp. 8–26.

[10] Michael Winikoff, "Assurance of agent systems: what role should formal verification play?," in *Specification and Verification of Multi-agent systems*, pp. 353–383. Springer, 2010.

[11] Hoang Tung Dinh and Tom Holvoet, "A framework for verifying autonomous robotic agents against environment assumptions," in *International Conference on Practical Applications of Agents and Multi-Agent Systems*. Springer, 2020, pp. 291–302.

[12] Marie Farrell, Matthew Bradbury, Michael Fisher, Louise A Dennis, Clare Dixon, Hu Yuan, and Carsten Maple, "Using threat analysis techniques to guide formal verification: A case study of cooperative awareness messages," in *International Conference on Software Engineering and Formal Methods*. Springer, 2019, pp. 471–490.

[13] Michael Winikoff, "Towards deriving verification properties," *arXiv preprint arXiv:1903.04159*, 2019.

[14] J Voas and K Schaffer, "Whatever happened to formal methods for security?," *Computer*, vol. 49, no. 8, pp. 70, 2016.

[15] William M McKeeman, "Differential testing for software," *Digital Technical Journal*, vol. 10, no. 1, pp. 100–107, 1998.

[16] Pradeep K Murukannaiah, Nirav Ajmeri, Catholijn M Jonker, and Munindar P Singh, "New foundations of ethical multiagent systems," in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 2020, pp. 1706–1710.

[17] N. Charron, S. Phillips, and S. L. Waslander, "De-noising of lidar point clouds corrupted by snowfall," in *2018 15th Conference on Computer and Robot Vision (CRV)*, 2018, pp. 254–261.

[18] Leilani H Gilpin, Jamie C Macbeth, and Evelyn Florentine, "Monitoring scene understanders with conceptual primitive decomposition and commonsense knowledge," *Advances in Cognitive Systems*, vol. 6, pp. 45–63, 2018.

[19] Robert M Lee, Michael J Assante, and Tim Conway, "German steel mill cyber attack," *Industrial Control Systems*, vol. 30, pp. 62, 2014.

[20] Lily Hay Newman, "The drone iran shot down was a $220m surveillance monster," *Wired.com*, 2019.

[21] Clayton Kunz, Chris Murphy, Hanumant Singh, Claire Pontbriand, Robert A Sohn, Sandipa Singh, Taichi Sato, Chris Roman, Ko-ichi Nakamura, Michael Jakuba, et al., "Toward extraplanetary under-ice exploration: Robotic steps in the arctic," *Journal of Field Robotics*, vol. 26, no. 4, pp. 411–429, 2009.

[22] Shizhe Zang, Ming Ding, David Smith, Paul Tyler, Thierry Rakotoarivelo, and Mohamed Ali Kaafar, "The impact of adverse weather conditions on autonomous vehicles: how rain, snow, fog, and hail affect the performance of a self-driving car," *IEEE vehicular technology magazine*, vol. 14, no. 2, pp. 103–111, 2019.

[23] AF Hixenbaugh, "Fault tree for safety," Tech. Rep., Boeing Co Seattle WA Support Systems Engineering, 1968.

[24] Bruce Schneier, "Attack trees," *Dr. Dobb's journal*, vol. 24, no. 12, pp. 21–29, 1999.

[25] Howard E Lambert, "Use of fault tree analysis for automotive reliability and safety analysis," *SAE transactions*, pp. 690–696, 2004.

[26] Gregory Falco, Arun Viswanathan, Carlos Caldera, and Howard Shrobe, "A master attack methodology for an ai-based automated attack planner for smart cities," *IEEE Access*, vol. 6, pp. 48360–48373, 2018.

FAULT TREE ANALYSIS AND RISK MITIGATION STRATEGIES FOR AUTONOMOUS SYSTEMS VIA STATISTICAL MODEL CHECKING

Ashkan Samadi, Marwan Ammar, and Otmane Ait Mohamed

Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
{ashkan.samadi, marwan.ammar, otmane.aitmohamed}@concordia.ca

ABSTRACT

In order to assess the reliability of autonomous systems, fault tree analysis (FTA) technique is used extensively. Most of the traditional FTA approaches are based on simulation and often require extensive computing capabilities. This paper proposes a formal FTA approach that can investigate the probability of failure of autonomous systems. The proposed methodology takes advantage of both FTA and statistical model checking (SMC). In order to illustrate the proposed approach, the sources of communication failure in a fleet of UAVs are analyzed. After detecting the most critical causes of communication failure, several redundant architectures are examined to assess their potentials to mitigate the risks of system failure. The results illustrate that all of the investigated architectures are capable of mitigating the probability of failure of the fleet of UAVs under studies.

Index Terms— fault tree analysis, statistical model checking, reliability analysis, fault mitigation

1. INTRODUCTION

In the past few years, the topic of safety analysis for autonomous systems has become increasingly challenging. One major attribute of an autonomous system is its ability to make decisions without human intervention. Therefore, it is necessary to ensure that the decision-making process of autonomous systems is error-free. In this context, many methodologies for fault analysis have been proposed. Among these approaches, fault trees (FTs) have been used extensively. A fault tree structure is a top-down graphical diagram that illustrates different events of lower levels that may result in the failure of the top-level event (TLE). In order to perform FTA, traditional techniques that are based on simulation can be utilized [1, 2]. However, simulation-based approaches require more resources, such as memory and processing time compared to analytical models [3] and because of their dependency on the sampling of the input space, an error can remain undiscovered unless all of the possible points are sampled. According to [4], several days might be needed to quantify all sequences of a large FT using the Monte Carlo method.

In recent years, formal-based approaches such as probabilistic model checking have been used for FTA. However, most of the probabilistic model checkers such as PRISM provide limited support for analyzing properties over time. In order to mitigate the aforementioned issue, statistical model checking (SMC) method [5, 6] can be used. SMC can analyze properties over time by simulating the model for a finite number of executions and using hypothesis testing to infer if the samples provide statistical evidence for or against a property. The ability of SMC to investigate the probability of system failure over time makes it feasible to determine the redundant architectures to reduce the probability of system failure according to the mission time of the system.

In this paper, a SMC based technique for FTA is proposed. Priced-timed automata (PTA) formalism is used to model the gates of the fault tree and to generate the fault tree model via parallel composition of the gates. The probability of system failure can be examined and the critical components of the FT can be detected by checking time-bounded reachability properties. Furthermore, several fault mitigation strategies are examined to investigate their potential to reduce the probability of system failure. We can decide on the best redundant configuration according to the intended mission time since our proposed methodology has the capability of assessing the probability of system failure over time. It is worth mentioning that the proposed probabilistic modeling of FTs is fully compatible with the concept of FT modularity [7]. It means that the FT of a large system can be constructed by connecting smaller sub trees which further reduces the chances of incurring into state-explosion.

In order to illustrate the proposed approach, the reliability of communication in a fleet of UAVs is analyzed using UPPAAL SMC tool. In our case-study, we reproduce the evaluation reported in [8]. In addition, more extensive comparative studies and predictive mitigation assessment can be performed. The rest of this paper is organized as follows: In section 2, the proposed methodology, including the fault tree modeling of systems is described. In section 3, the results of the experiments are reported, and in section 4, our conclusions are presented.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

2. PROPOSED METHODOLOGY

In this section, our proposed approach is introduced as depicted in Fig. 1. First, a library with the PTA models of all basic gates of the FT is generated. Together with the predefined basic events and their probabilities, these gates are used to model the components of the system. The FT is then obtained by the parallel composition of all its components. Subsequently, the constructed FT model is assessed to examine the probability of system failure. If the failure probability is greater than the threshold specified in the mission requirements, then a Component Contribution Investigation (CCI) is performed. CCI investigates the contribution of each component to the probability of system failure. This is done by verifying properties that determine the probability of system failure when failure occurs in only one specific component of the fault tree. For example, in Fig. 5, by checking 4 properties that each one specifies the occurrence of top-level event and one second-level event (e.g. software failure) and the non-occurence of other second-level events (SLEs), the contribution of SLEs to system failure can be investigated. Then, the probability of system failure obtained by verifying each property is stored in a table. By examining the table, the most critical SLE that causes higher probability of system failure can be detected. The same analysis is performed on the components in lower levels until the critical components of the FT is detected from the table. In the next step, the proposed methodology evaluates the effects of applying several redundant configurations to the critical components (i.e., higher weight contribution to system failure). If the probability of system failure remains above the threshold, the process will be repeated until either the probability of system failure falls below the target threshold or the threshold is deemed unreachable.

2.1. Fault Trees Modeling

In this work, UPPAAL-SMC is utilized to model the probabilistic behavior of the FT gates and the events as PTA, adopted from [9]. The aforementioned tool performs statistical model checking to verify formal specifications of stochastic systems [10].

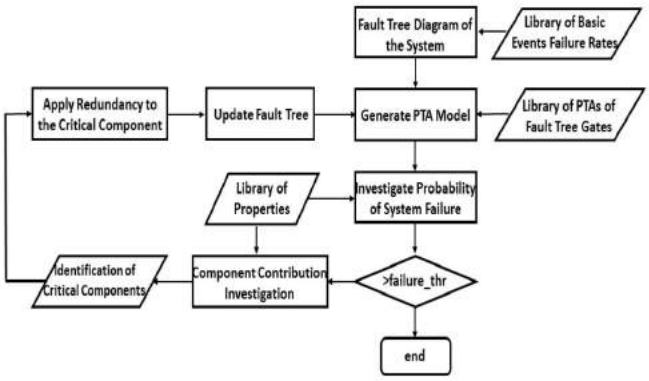


Fig. 1. Main steps of the proposed approach.

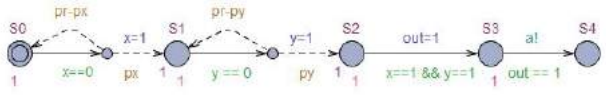


Fig. 2. CSP Gate

In Fig. 2, a possible configuration of a *cold spare (CSP)* gate with one primary input and one alternate (spare) input is shown as an example. At first, the primary input is in active mode, and the spare input is turned off. When the primary input fails, the alternate input is switched to replace it. For this model, no activation delay is assumed and the switcher is not taken into account. We assume that the probability of failure of the primary event as *px* and the alternate event as *py*. The initial state of *S0* represents the absence of faults. If the primary event fails, variable *x* becomes one and then, the automaton moves to state *S1*. If the alternate event fails, the automaton can then move to state *S2* with probability *py*. Afterward, if *x==1* and *y==1* (both the primary input and the alternate input fail), the variable *out* is set to 1. Next, the broadcast is emitted through broadcast channel *a*. The complete model of the FT is produced using a network of PTA models that are connected via broadcast channels.

2.2. TMR Modeling

Triple modular redundancy is an architectural pattern that is employed widely for component redundancy [11, 12, 13]. In TMR configurations a particular component is triplicated and one or more majority voters are used to decide the output by computing the majority outcome among the inputs. Because applying the physical TMR configuration to the complete system may be expensive in terms of monetary and space constrains, it is important to determine the most efficient places to implement it.

In our proposed methodology, the components that are most likely to fail in the studied system are selected for applying different TMR configurations. In [14], several TMR arrangements including 2-stage TMR configurations with 1 and 3 voters are investigated. In our experiments, we analyze 2-stage and 3-stage arrangements with TMR configurations incorporating 1 and 3 voters. In Fig. 3, an example of a chain of 3-voter TMR configurations is illustrated. The investigated configurations are adopted from [15].

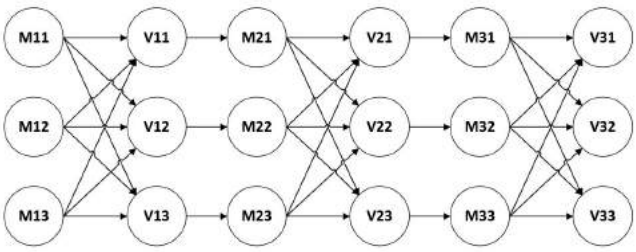


Fig. 3. Example of a 3 stage TMR arrangement.

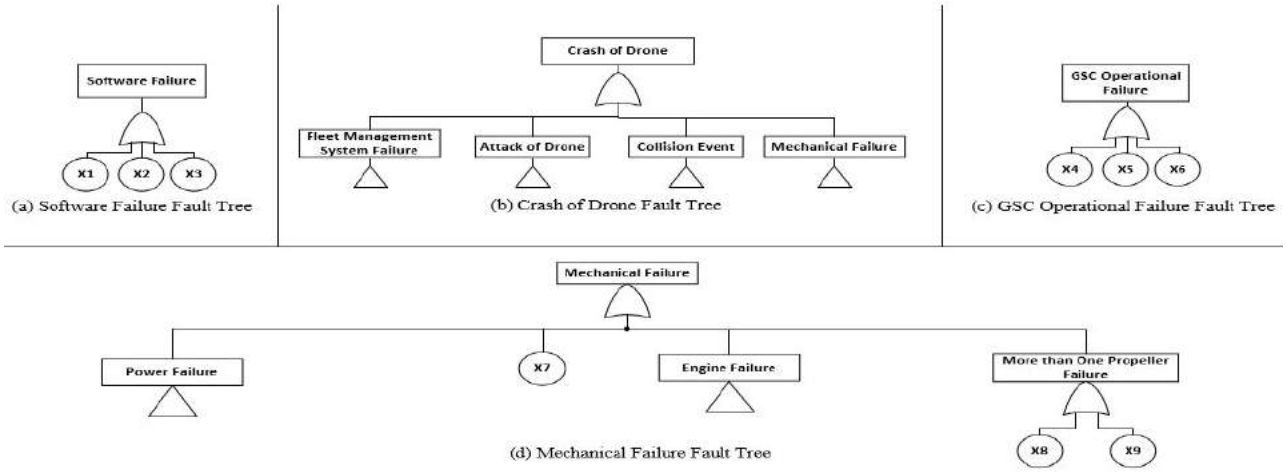


Fig. 4. FTs of several intermediate events.

3. EXPERIMENTAL RESULTS

In this section, the results of the analyses on the FTs in Fig. 4 and 5 are presented. After detecting the most critical components, several redundant configurations are investigated to determine which architecture offers optimal failure mitigation. The experiments have been conducted on a machine with an Intel Core I7-6700HQ CPU and 8 GB of RAM.

The FTs in this experiment and the mean failure rates of the basic events are adopted from [8]. Their results illustrate that *crash of drone* is the most critical second-level event (SLE) among the intermediate events that are connected to the top event. According to [8], there are several elements that can cause communication failure in a fleet of drones. In Fig. 5, the four intermediate events that affect the communication in the fleet are shown. With respect to the *crash of drone*, certain factors such as *mechanical failure* and *fleet management system failure* are mentioned. In the fault tree of Fig. 4(b), the four intermediate events that can potentially result in the *crash of drone* are represented. In Fig. 4(d), the sources of *mechanical failure* are illustrated. Two other SLEs, namely *software failure* and *GSC operational failure* are illustrated in Fig. 4(a) and Fig. 4(c), respectively. The complete FT of the system is not included in this paper due to space constraints. Once the complete FT is composed, the cumulative probability distribution (CPD) of the top event failure can be plotted in an interval expressed in time-units using the following time-bounded reachability property:

$$Pr[<= 20](<> CommunicationFailure) \qquad (1)$$

This property specifies the probability confidence interval that *CommunicationFailure* will be asserted (failure occurs) eventually within 20 time-units. The time window of 20 time-units has been chosen because according to our experiments the CPD normalizes within 20 time-units. It is worth mentioning that the elapsed time and the memory usage to check the aforementioned property are 0.083s and 37.4 MB, respectively, showing that our methodology is extremely efficient and very scalable.

In order to study the effect of each second-level event separately, query (1) is modified to specify the probability that the top-event as well as one of the 4 SLEs are asserted (the failure occurs) but the other 3 SLEs do not occur. As an example, with checking the reachability property of $Pr[<= 20](<> CommunicationFailure == 1 \; and \; CrashDrone == 1 \; and \; SoftwareFailure == 0 \; and \; GSCFailure == 0 \; and \; LossofConnection == 0)$ the contribution of *crash of drone* to the CPD of top event failure within 20 time-units can be plotted. After using 4 similar variants of the aforementioned query (each one considering the occurrence of the top event as well as only one SLE and non-occurrence of the other 3 SLEs), the five curves of Fig. 6 are plotted. The curve with the highest value, is the CPD of TLE failure. The other 4 curves illustrate the contribution of each second-level event to the CPD of *communication failure* in the interval of 20 time-units. All of the curves start with similar values. However, the contribution of *crash of drone* to the CPD of TLE failure grows at a higher rate than the other SLEs and after 12 units of time the curves become stable. It is observed that the probability of *communication failure* is influenced by *crash of drone* more than the other SLEs and therefore it is the most critical second-level event.

Fig. 5. FT of communication failure in a fleet of UAVs.

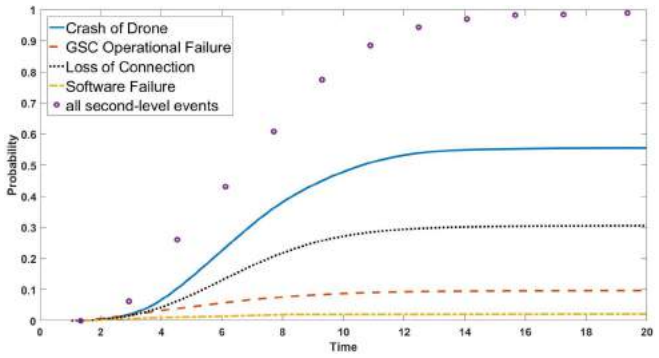


Fig. 6. Contribution of SLEs to the CPD of system failure.

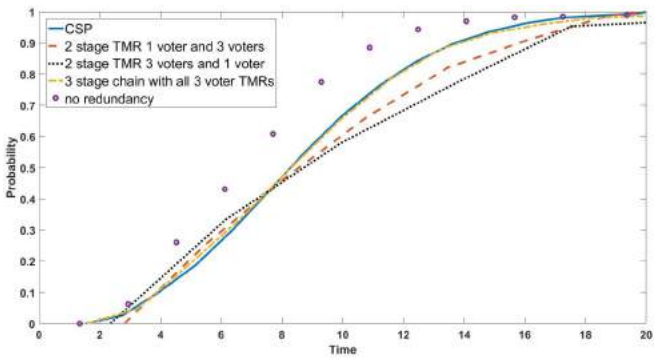


Fig. 7. Effect of redundancy on the CPD of system failure.

After specifying the most critical SLE, we continue to analyze more components of the FT to detect the critical ones. For this purpose, the same analysis which was conducted for the top event failure is repeated for *crash of drone*. With the property of $Pr[<= 20](<> CrashDrone)$ the CPD of *crash of drone* event can be plotted within 20 time-units. Furthermore, the aforementioned query is modified to consider the occurrence of fault in just one of the 4 events that are connected to *crash of drone*. Therefore, 4 other curves can be plotted to illustrate the contribution of each intermediate event to the CPD of *CrashDrone*. The results show that the curve of the CPD of *crash of drone* is more affected by the intermediate event of *mechanical failure*. Thus, *mechanical failure* is the most critical intermediate event which is connected to *crash of drone*. Afterward, similar to the previous experiments, we analyze the contribution of the lower-level components to the CPD of *mechanical failure*. The results illustrate that *more than one propeller failure* is the most critical basic component of the FT.

After detecting the most critical component, several redundant architectures based on multiple stage chains of TMRs as well as the cold spare are examined. In order to investigate the effect of the cold spare, the critical component is duplicated and one of the two instances is used as the primary input of the CSP and the other instance as the spare input. The results in Fig. 7 depicts the CPD of system failure considering the effects of applying different redundant architectures to the critical component. It is observed that the failure rate of the TLE decreases after applying the redundancies based on the cold spare, 2 stage and 3 stage chains of TMRs with 1 voter and 3 voters. Although the difference among the effects of applying the redundancies might seem negligible, it must be taken into account that the failure rates of basic events are in the range of 1.0E-03 or even lower. Since such small failure rates are considered for the experiments, even small differences can be significantly impactful, especially over long mission times. In Fig. 7, it can be observed that the chain of TMRs with the first stage as a 3 voter configuration and the second stage as a 1 voter configuration results in better fault tolerance after almost 8 time units of system operation. In the interval from the start of the operation of the system until almost 4 time units the chain of TMRs with the first stage as a 1 voter configuration and the second stage as a 3 voter configuration is more effective and after this instance of time until almost 8 time units cold spare can reduce the risk of system failure more than the other investigated redundant architectures. Thus, it is necessary to consider the expected mission time of the system when selecting the type of redundancy. It is worth mentioning the TMR configuration with 1 voter is more advantageous than the TMR with 3 voters in terms of area and power and any constraints on the aforementioned metrics can also be impactful for a specific system.

4. CONCLUSION

In this paper, a fault tree analysis technique based on statistical model checking is proposed. The proposed approach can investigate the probability of failure of autonomous systems over time which is not often feasible using other model checkers. Our approach can determine the contribution of each lower-level event to the probability of higher-level events using properties without the need to make any changes to the fault tree model. In order to depict the proposed methodology, the sources of communication failure in a fleet of drones are investigated. Subsequently, the critical components that may cause system failure are detected. In addition, several redundant configurations are tested to examine their capabilities of fault mitigation. According to the results, the chain of TMRs with the first stage as a 1-voter configuration and the second stage as 3-voter outperforms the other examined redundancies before 4 units of mission time.

5. REFERENCES

[1] K Durga Rao, V Gopika, VVS Sanyasi Rao, HS Kushwaha, Ajit Kumar Verma, and Ajit Srividya, "Dynamic fault tree analysis using monte carlo simulation in prob-

abilistic safety assessment," *Reliability Engineering & System Safety*, vol. 94, no. 4, pp. 872–883, 2009.

[2] Gabriele Manno, Ferdinando Chiacchio, Lucio Compagno, Diego D'Urso, and Natalia Trapani, "Matcarlore: An integrated ft and monte carlo simulink tool for the reliability assessment of dynamic fault tree," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10334–10342, 2012.

[3] Sohag Kabir, Koorosh Aslansefat, Ioannis Sorokos, Yiannis Papadopoulos, and Savas Konur, "A hybrid modular approach for dynamic fault tree analysis," *IEEE Access*, vol. 8, pp. 97175–97188, 2020.

[4] Sang Hoon Han, "A top-down iteration algorithm for monte carlo method for probability estimation of a fault tree with circular logic," *Nuclear Engineering and Technology*, vol. 50, no. 6, pp. 854–859, 2018.

[5] Koushik Sen, Mahesh Viswanathan, and Gul Agha, "Statistical model checking of black-box probabilistic systems," in *International Conference on Computer Aided Verification*. Springer, 2004, pp. 202–215.

[6] Hakan L Younes, "Verification and planning for stochastic processes with asynchronous events," Tech. Rep., CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE, 2005.

[7] William E Vesely, Francine F Goldberg, Norman H Roberts, and David F Haasl, "Fault tree handbook," Tech. Rep., Nuclear Regulatory Commission Washington DC, 1981.

[8] Rana Abdallah, Raed Kouta, Charles Sarraf, Jaafar Gaber, and Maxime Wack, "Fault tree analysis for the communication of a fleet formation flight of uavs," in *2017 2nd International Conference on System Reliability and Safety (ICSRS)*. IEEE, 2017, pp. 202–206.

[9] Marwan Ammar, Ghaith Bany Hamad, Otmane Ait Mohamed, and Yvon Savaria, "Towards an accurate probabilistic modeling and statistical analysis of temporal faults via temporal dynamic fault-trees (tdfts)," *IEEE Access*, vol. 7, pp. 29264–29276, 2019.

[10] Alexandre David, Kim G Larsen, Axel Legay, Marius Mikučionis, and Danny Bøgsted Poulsen, "Uppaal smc tutorial," *International Journal on Software Tools for Technology Transfer*, vol. 17, no. 4, pp. 397–415, 2015.

[11] Peter A Lee and Tom Anderson, *Fault tolerance, principles and practice*, vol. 3, Springer Verlag, 1990.

[12] Michele Favalli and Cecilia Metra, "Tmr voting in the presence of crosstalk faults at the voter inputs," *IEEE Transactions on Reliability*, vol. 53, no. 3, pp. 342–348, 2004.

[13] Darshan D Thaker, Rajeevan Amirtharajah, Francois Impens, Isaac L Chuang, and Frederic T Chong, "Recursive tmr: Scaling fault tolerance in the nanoscale era," *IEEE Design & Test of Computers*, vol. 22, no. 4, pp. 298–305, 2005.

[14] Matthew J Cannon, Andrew M Keller, Corbin A Thurlow, Andrés Pérez-Celis, and Michael J Wirthlin, "Improving the reliability of tmr with nontriplicated i/o on sram fpgas," *IEEE Transactions on Nuclear Science*, vol. 67, no. 1, pp. 312–320, 2019.

[15] Marwan Ammar, Ghaith Bany Hamad, Otmane Ait Mohamed, and Yvon Savaria, "Efficient probabilistic fault tree analysis of safety critical systems via probabilistic model checking," in *2016 Forum on Specification and Design Languages (FDL)*. IEEE, 2016, pp. 1–8.

TOWARDS THREE-DIMENSIONAL ACTIVE INCOHERENT MILLIMETER-WAVE IMAGING

Stavros Vakalis and Jeffrey A. Nanzer

Electrical and Computer Engineering, Michigan State University

ABSTRACT

Active incoherent millimeter-wave (AIM) imaging is a new technique that combines aspects of passive millimeter-wave imaging and noise radar to obtain high-speed imagery. Using an interferometric receiving array combined with small set of uncorrelated noise transmitters, measurements of the Fourier transform domain of the scene can be rapidly obtained, and scene images can be generated quickly via two-dimensional inverse Fourier transform. Previously, AIM imaging provided two-dimensional reconstructions of the scene. In this work we explore the use of active millimeter-wave imaging for automotive sensing by investigating array feasible layouts for automobiles, and a new technique to impart range estimation to obtain three-dimensional imaging information.

Index Terms— Millimeter-wave imaging, high-speed imaging, noise radar, automotive radar

1. INTRODUCTION

Millimeter-wave imaging has benefits for a wide range of applications, including security sensing [1], contraband detection [2], medical imaging [3], and non-destructive testing [4], among others. Recently, millimeter-wave imaging for automotive radar has become of significant interest due to the rapidly evolving field of vehicle autonomy [5, 6, 7, 8]. Millimeter-wave imaging holds significant potential for automotive applications due to the ability of millimeter-wave radiation to propagate through obscurants like fog, smoke, snow, and light rain with little to no impact, while at the same time maintaining good imaging resolution due to the short wavelengths of millimeter-wave signals [1].

Millimeter-wave automotive sensing uses active transmission of signals, rather than passive techniques, to ensure sufficient sensing range and operational speed. Active imaging at millimeter-wave frequencies traditionally relies on the transmission and reception of a coherent radar signal combined with a narrow beam steered either mechanically or electrically [9]. Mechanical imagers tend to be bulky and slow [10], limiting their feasibility in automotive applications, while electrically-scanned systems require a significant number of electrical components and a large aperture to generate high-resolution imagery, driving up cost and power consumption. Multiple-input and multiple-output (MIMO) techniques combined with frequency-modulated continuous-wave (FMCW) radar waveforms can achieve imaging without analog beamforming [11], however MIMO FMCW radar systems require complex synchronization between the receive and transmit array elements, and also entail a significant amount of additional processing since each transmit signal is processed orthogonally on each receiving element [12].

In this work we introduce a new concept for three-dimensional automotive sensing that builds on a recently developed active incoherent millimeter-wave (AIM) imaging technique. In contrast to traditional millimeter-wave techniques, AIM imaging combines the transmission of noise signals with a sparse interferometric receiving aperture to obtain high-resolution two-dimensional millimeter-wave imagery at video rates, without scanning [13, 14]. Here we explore a new concept that imparts a coarse time synchronization of the transmitted noise waveforms to obtain range resolution and thus imaging in three dimensions. We investigate two array layouts commensurate with implementation on a vehicle facade at 77 GHz and investigate the three-dimensional response of the system with a wideband pulsed noise waveform.

2. ACTIVE INCOHERENT MILLIMETER-WAVE (AIM) IMAGING

Interferometric imaging was first developed in radio astronomy to observe the radiation from stars and other stellar objects using sparse arrays to sample the spatial Fourier transform of the signals in the array field of view [15], but has since been applied to satellite remote sensing and security sensing. Interferometric arrays sample the Fourier transform of the scene intensity, instead of sampling the scene intensity directly like other imaging modalities. Interferometric imaging systems can use sparse antenna arrays requiring significantly fewer antenna elements than traditional phased arrays, while also maintaining a tolerance to element failures [16, 17]. Interferometric arrays have traditionally been passive systems, capturing thermal emissions from the scene. Interferometric imaging necessitates that the received signals are spatially

This work was supported by the National Science Foundation under Grant 1708820.

and temporally incoherent to reconstruct the scene from the Fourier domain samples [15], and thermal radiation naturally satisfies this constraint. However, thermal radiation power is generally exceedingly small at millimeter-wave frequencies, requiring highly sensitive receivers with high gain and long integration times, limiting the practicability of the technique for automotive applications, where fast image reconstruction time and low cost is important.

In AIM imaging, we combine the benefits of active and passive millimeter-wave imaging. Active millimeter-wave systems operate with significantly higher signal-to-noise ratio (SNR) due to the transmission of signals, and therefore do not require receivers with high sensitivity. Passive interferometric millimeter-wave systems employ very sparse antenna arrays and, furthermore, generate imagery in a staring format, without beamsteering. We previously demonstrated AIM imaging at microwave and millimeter-wave frequencies using multiple noise transmitters to effectively mimic the properties of thermal radiation and support Fourier-domain image reconstruction [13, 14].

Interferometric antenna arrays, whether used for passive or AIM imaging, capture samples of the scene visibility $\mathcal{V}(u,v)$, which is the spatial Fourier transform of the scene where (u,v) are spatial frequencies. Samples of the visibility are obtained via cross-correlation of the signals between pairs of antennas, yielding a sampling function $S(u,v)$, where the sampled spatial frequencies are defined by the electrical separation and rotation angle of the antenna pairs. The product of the sampling function $S(u,v)$ and $\mathcal{V}(u,v)$ is referred to as the sampled visibility $\mathcal{V}_s(u,v)$, from which the image intensity I_r can be reconstructed through an inverse Fourier transform

$$I_r(\alpha,\beta) = \iint_{-\infty}^{\infty} \mathcal{V}_s(u,v) e^{-j2\pi(u\alpha+v\beta)} du dv \qquad (1)$$

where α and β are the direction cosines in the azimuth and elevation plane. The spatial response of an interferometric array can be characterized by the point spread function, which is given by

$$PSF = \mathcal{F}^{-1}\Big[S(u,v)\Big]. \qquad (2)$$

In incoherent imaging the point spread function typically refers to the squared magnitude $|PSF|^2$.

3. ANALYSIS OF AIM IMAGING FOR 77 GHZ AUTOMOTIVE RADAR

We consider the imaging performance of two 24-element 77 GHz array layouts for potential implementation in the facade of a vehicle: a randomized aperture that has the same azimuth and elevation resolution and a randomized aperture using the same number of elements that increases resolution in the azimuth plane at the expense of resolution in the elevation plane. Wider inter-element spacing can lead to larger

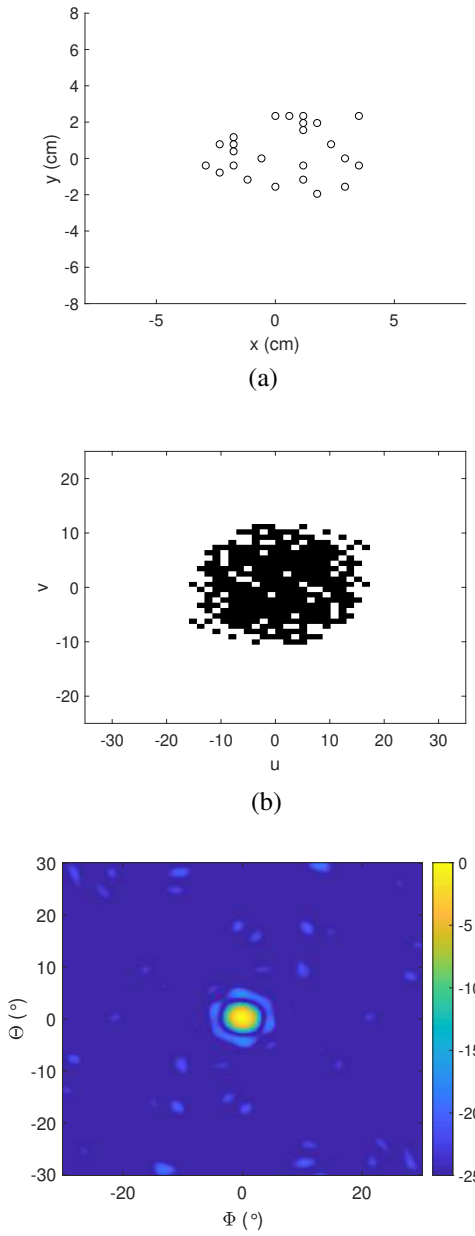


Fig. 1. (a) Randomized 24-element antenna array. The minimum spacing in both the horizontal and vertical dimension is 1.5 λ. (b) Sampling function of the random 24-element aperture. (c) Point spread function of the randomized aperture as a function of the azimuth and elevation angles Φ and Θ.

electrical aperture maximum dimensions, and therefore improved resolution, however this also introduces image ambiguities, limiting the effective field-of-view of the imager. The half-angle unambiguous field of view of an interferometric imager with element spacing d_x and d_y across the horizon-

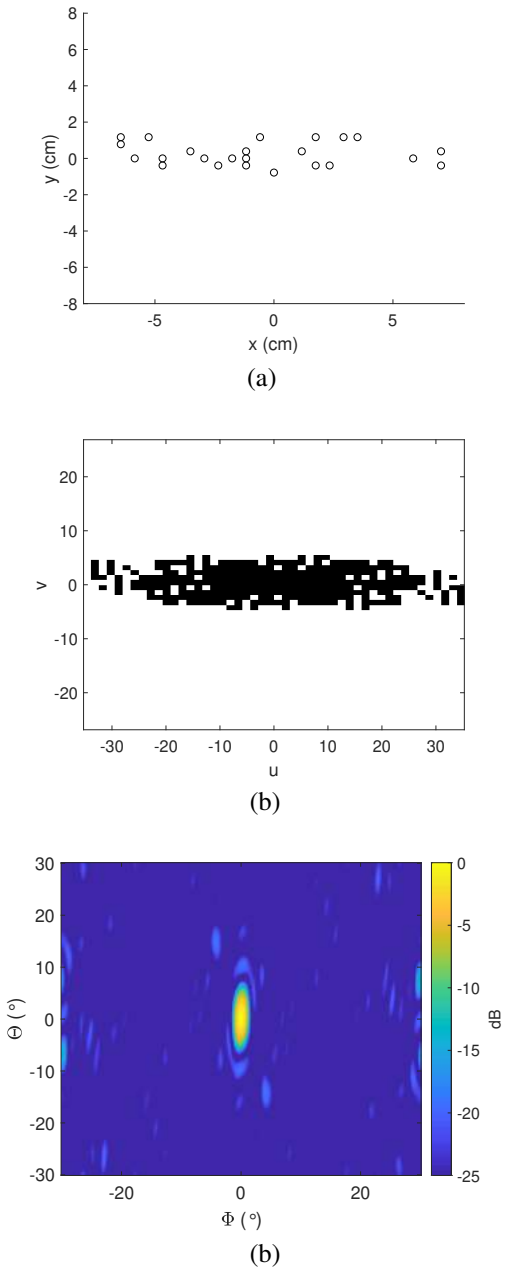


Fig. 2. (a) Random 24-element antenna array. The minimum spacing in both the horizontal and vertical dimension is 1.5 λ. (b) Sampling function of the random 24-element aperture. The sampling function is significantly wider in the u dimension. (c) Point spread function of the randomized aperture as a function of the azimuth and elevation angles Φ and Θ. The beamwidth is significantly larger in the elevation plane than in the azimuth plane, due to the smaller aperture electrical dimensions.

tal and vertical axes can be expressed for the two direction cosines α and β as

$$FOV_{\frac{\alpha}{2},\frac{\beta}{2}} = \frac{\lambda}{2 \cdot d_{x,y}} \quad . \tag{3}$$

The 24-element randomized antenna array with equal azimuth and elevation resolution is shown in Fig. 1(a). The minimum spacing between two antenna elements is 0.58 cm (1.5λ), which corresponds to an ambiguous field of view of approximately $\pm 34°$ in the azimuth (Φ) and elevation (Θ) angles. Due to its less stringent layout, random arrays allow more flexibility for integration into different vehicle facades. The sampling function $S(u, v)$ of the array is shown in 1(b). The PSF of the array is shown in 1(c), which shows a resolution of 7 degrees in both azimuth and elevation dimensions. At a distance of 20 m from the aperture, the resolution is 2.42 m in the azimuth (cross-range) dimension. While this array has equivalent elevation and azimuth resolution, for automotive applications azimuth resolution is generally more important, thus we designed a second randomized aperture using the same number of elements, but spanning a wider horizontal space and a narrower vertical space. Since the resolution is inversely proportional to the maximum aperture size, this serves to improve the resolution in the azimuth dimension at the expense of resolution in the elevation dimension. In this case, the array was four times wider in the horizontal dimension that the vertical dimension. The second randomized 24-element antenna array with 1.5λ minimum spacing in both the x and y dimension is shown in Fig. 2(a). The sampling function $S(u, v)$ of the array is shown in 2(b). The point spread function can be seen in Fig. 2(c). The array maximum vertical dimension has decreased in this array, and therefore the resolution has become larger in the elevation plane and equal with 15 degrees, however the resolution has improved in the azimuth plane to 3.3 degrees, which is of greater importance. At a distance of 20 m from the aperture, the resolution is 1.16 m in the azimuth (cross-range) dimension. Apertures with wider horizontal coverage or more antenna elements could easily be designed to further improve the resolution.

4. 3-D AIM IMAGING

Interferometric imaging, both passive and active, do not inherently provide for a mechanism to obtain range information. For automotive applications, down-range measurements should generally be combined with cross-range measurements for accurate environmental sensing. Techniques using volumetric arrays or near-field processing have been explored [18, 19], however their requirements are not practical for automotive applications. Traditional coherent processing techniques, such as matched filtering, are not possible in passive interferometry since no transmit signal is used. In AIM, the specifics of the transmit signals need not be known to simplify the hardware requirements; simple noise emitters

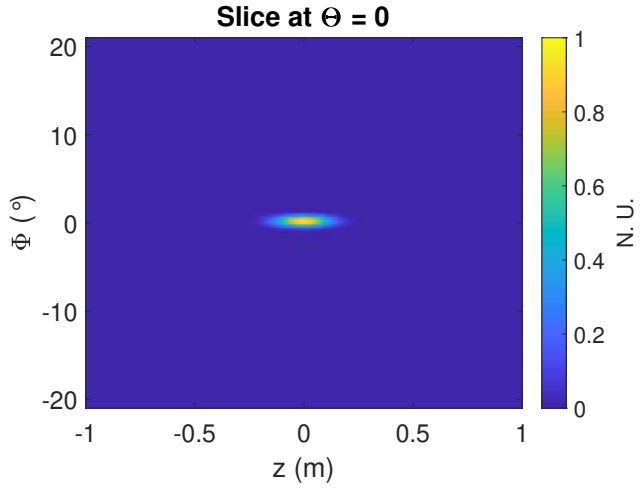


Fig. 3. Two-dimensional $\Phi - z$ slice of the $PSF(\Phi, \Theta, z)$ from the randomized aperture in Fig. 2(a). The bandwidth of the pulse is 200 MHz.

can be used as long as their statistics are known, precluding matched filtering.

In this work we explore a new approach of pulse modulating the transmitted noise signals to obtain down-range information. We control the timing of the transmitted signal envelopes, which can be done with relatively simple coordination and does not impose spatial coherence, thereby preserving the cross-range incoherence. The result effectively generates two-dimensional interferometric images sequentially in time, each time segment representing a different range bin. We analyze the PSF of the system using a Gaussian pulse on the transmitted noise signals. The frequency-domain signal on each transmitter can be written as

$$S_i(f) = e^{\frac{-(f-f_c)^2}{4\delta f^2}} N_i(f) \tag{4}$$

where $N_i(f)$ is the spectrum of the ith wideband Gaussian noise signal, f_c is the carrier frequency and δf is the pulse bandwidth. The range resolution Δz along the z dimension can be approximated by the full width at half maximum of a squared Gaussian pulse as

$$\Delta z \approx \frac{2.3548\, c}{\sqrt{2}\pi\delta f}. \tag{5}$$

For δf = 200 MHz, $\Delta z \approx 20$ cm. For the three-dimensional simulations we assume that the carrier frequency of the imaging system is 77 GHz and the bandwidth δf is 200 MHz. A slice $\Phi - z$ of the point spread function $PSF(\Phi, \Theta, z)$ for the random array is shown in Fig. 3, demonstrating the down-range and cross-range resolution capabilities.

5. CONCLUSION

We investigated the point spread function of a new approach to achieve three-dimensional active incoherent millimeter-wave imaging. Array layouts were considered that matched automotive radar frequencies and physical sizes that are commensurate with potential application in an automotive facade. We investigated the use of pulse modulation on incoherent transmitted signals to add three-dimensional measurement capability to a previously demonstrated two-dimensional measurement system. Future efforts may build on this work to implement fast, robust environmental sensing for automotive applications.

6. REFERENCES

[1] J. A. Nanzer, *Microwave and Millimeter-Wave Remote Sensing for Security Applications*, Artech House, 2012.

[2] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 9, pp. 1581–1592, Sep 2001.

[3] N. K. Nikolova, "Microwave imaging for breast cancer," *IEEE Microw. Mag.*, vol. 12, no. 7, pp. 78–94, Dec 2011.

[4] F. Zidane, J. Lanteri, J. Marot, L. Brochier, N. Joachimowicz, H. Roussel, and C. Migliaccio, "Nondestructive control of fruit quality via millimeter waves and classification techniques: Investigations in the automated health monitoring of fruits," *IEEE Antennas Propag. Mag.*, vol. 62, no. 5, pp. 43–54, 2020.

[5] Igal Bilik, Shahar Villeval, Daniel Brodeski, Haim Ringel, Oren Longman, Piyali Goswami, Chethan Y. B. Kumar, Sandeep Rao, Pramod Swami, Anshu Jain, Anil Kumar, Shankar Ram, Kedar Chitnis, Yashwant Dutt, Aish Dubey, and Stanley Liu, "Automotive multi-mode cascaded radar data processing embedded system," in *2018 IEEE Radar Conference (RadarConf18)*, 2018, pp. 0372–0376.

[6] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 ghz frequency band," *IEEE Trans. Microw. Theory Tech.*, vol. 60, no. 3, pp. 845–860, 2012.

[7] J. Dickmann, J. Klappstein, M. Hahn, N. Appenrodt, H. Bloecher, K. Werber, and A. Sailer, "Automotive radar the key technology for autonomous driving: From detection and ranging to environmental understanding," in *2016 IEEE Radar Conference (RadarConf)*, 2016, pp. 1–6.

[8] Martin Stolz, Mingkang Li, Zhaofei Feng, Martin Kunert, and Wolfgang Menzel, "High resolution automotive radar data clustering with novel cluster method," in *2018 IEEE Radar Conference (RadarConf18)*, 2018, pp. 0164–0168.

[9] B. Ku, P. Schmalenberg, O. Inac, O. D. Gurbuz, J. S. Lee, K. Shiozaki, and G. M. Rebeiz, "A 77–81-ghz 16-element phased-array receiver with $\pm 50^{\circ}$ beam scanning for advanced automotive radars," *IEEE Trans. Microw. Theory Tech.*, vol. 62, no. 11, pp. 2823–2832, 2014.

[10] Josiah Wayl Smith, Muhammet Emin Yanik, and Murat Torlak, "Near-field mimo-isar millimeter-wave imaging," in *2020 IEEE Radar Conference (RadarConf20)*, 2020, pp. 1–6.

[11] R. Feger, C. Wagner, S. Schuster, S. Scheiblhofer, H. Jager, and A. Stelzer, "A 77-ghz fmcw mimo radar based on an sige single-chip transceiver," *IEEE Trans. Microw. Theory Tech.*, vol. 57, no. 5, pp. 1020–1035, May 2009.

[12] Francesco Belfiori, Wim van Rossum, and Peter Hoogeboom, "Random transmission scheme approach for a fmcw tdma coherent mimo radar," in *2012 IEEE Radar Conference*, 2012, pp. 0178–0183.

[13] S. Vakalis and J. A. Nanzer, "Microwave imaging using noise signals," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 12, pp. 5842–5851, Dec 2018.

[14] S. Vakalis, L. Gong, Y. He, J. Papapolymerou, and J. A. Nanzer, "Experimental demonstration and calibration of a 16-element active incoherent millimeter-wave imaging array," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 9, pp. 3804–3813, 2020.

[15] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy*, John Wiley and Sons, 2001.

[16] S. Vakalis and J. A. Nanzer, "Analysis of array sparsity in active incoherent microwave imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 1, pp. 57–61, Jan 2020.

[17] S. Vakalis and J. A. Nanzer, "Analysis of element failures in active incoherent microwave imaging arrays using noise signals," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 2, pp. 161–163, Feb 2019.

[18] Joseph Rosen and Amnon Yariv, "Three-dimensional imaging of random radiation sources," *Opt. Lett.*, vol. 21, no. 14, pp. 1011–1013, Jul 1996.

[19] N. A. Salmon, "3-d radiometric aperture synthesis imaging," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 11, pp. 3579–3587, 2015.

An Autonomous Semantic Learning Methodology for Fake News Recognition

Yingxu Wang, *Fellow, IEEE*, and James Y. Xu, *Graduate Student Member, IEEE*

International Institute of Cognitive Informatics and Cognitive Computing (ICICC)
Department of Electrical and Software Engineering
Schulich School of Engineering and Hotchkiss Brain Institute
University of Calgary, Canada
Emails: *yingxu*@ucalgary.ca and *yifan.xu1@ucalgary.ca*

Abstract — A persistent challenge to AI theories and technologies is fake news recognition which demands not only syntactic analyses of language expressions, but also their semantics comprehension. This work presents an autonomous system for fake news recognition based on a novel approach of machine semantic learning. A training-free machine learning algorithm of Differential Sentence Semantic Analyses (DSSA) is designed and implemented for fake news detection. A large set of 876 experiments randomly selected from DataCup'19 has demonstrated a level of 70.4% accuracy that outperforms the traditional data-driven neural network technologies normally projected at the accuracy level of 55.0%. The DSSA methodology paves a way towards autonomous, training-free, and real-time trustworthy technologies for machine knowledge learning and semantics composition.

Keywords — Fake news, autonomous systems, knowledge learning, semantic comprehension, intelligent mathematics, cognitive systems

I. INTRODUCTION

Fake news recognition is a persistent challenge to AI theories and technologies because it demands both syntactic and semantic learning and analyses [1-7]. A piece of news is considered fake when its semantics is inconsistent to the contextual facts and their syntactic constraints. Therefore, fake news may be classified as syntactically, semantically, and/or sequentially inconsistencies in natural language expressions. It is recognized that the field of fake news recognition is dependent on the 6th form of machine learning as developed in our lab known as knowledge learning [8] which is beyond traditional machine learning capabilities on object identification, cluster classification, pattern recognition, functional regression, and behavioral generation [9-11].

Classical approaches to fake news recognition are based on news properties including source reliability, authority, acceptable rates, citations, causality, followers, and community support [3-6, 8]. Many recent technologies adopt a regression classifier to label a sentence as true or false [1, 7, 12], while they may not be able to distinguish simple inconsistent expressions such as "He is a student" vs. "He had been a student" and which is fake? This is due to the universe of discourse and the rigorous semantic pattern of fake news have not been formally studied [13-14]. In the literature, many technologies have been proposed for fake news recognition. A common approach focuses on the analysis and extraction of text features by neural networks [4, 12]. However, since there is no rigorous pattern and characteristics of fake news, the trueness and trustworthiness of regression results are not reliable and normally closer to a random guess at the level of 50+% accuracy. For instance, the best results of DataCup'19 [2] was constrained approximately 55.0%, because there is no matured literature training technology for neural networks in natural languages. The state-of-the-art may be compared with human learning mechanisms where the latter are dependent on semantic learning, while the former are driven by data regressions.

It is recognized that the 6th form and most important machine learning is knowledge learning. In machine knowledge learning, semantics comprehension deals with how meanings of a sentence in a language are conveyed and comprehended. Semantical learning is the frontier of cognitive linguistics that studies the interpretation of meanings of words and sentences as a mapping from the set of unknown words to known ones and their compositions [14, 15]. Formal semantics are studied in classic linguistics [15, 16] and cognitive linguistics [13, 17-19].

This paper presents an autonomous system for fake news recognition based on machine learning for semantic comprehension in the context of the DataCup'19 competition for fake news detection organized in Canada [2]. In the remainder of this paper, the mathematical model of fake news is introduced in Section II driven by a paradigm of intelligent mathematics known as *semantic algebra* [14]. Then, a training-free learning algorithm of differential semantic analyses is formally described in Section III for fake news detection. It leads to the implementation of the algorithm that are validated by a large set of experiments for fake news recognition that outperforms traditional technologies.

II. THE FORMAL MODEL OF FAKE NEWS

Because fake news is an inconsistent statement of an event against its underpinning facts, its detection is highly dependent

on a formal model of linguistic semantics for machine-enabled *semantic comprehensions* of natural language expressions in both claims and associated facts. The semantics of a natural language can be classified into three categories [14] known as those of *entities* (nouns and noun phrases), *behaviors* (verbs and verb phrases), and *modifiers* (adjectives, adverbs, and related phrases). Complex semantics beyond those of words can be coherently aggregated from the bottom up. Semantics can also be classified into the categories of *to-be* and *to-do* semantics according to semantic algebra [14].

2.1 The Semantic Model of Natural Language Expressions

Semantics is the composed meaning of language expressions and perceptions at the levels of words, phrase, sentence, paragraph, and essay. The basic unit of formal semantics is a concept as a formal model of words in a natural language [10]. The semantic models of fake news are focused on those of the *to-be* (noun phrase) and *to-do* (verb) structures, particularly the latter.

Definition 1. The *semantics of an entity E*, $\Theta(E)$, is embodied by a *formal concept* C_E, denoted by $|=:$

$$\Theta(E) \triangleq \Theta(E\ \_\ C_E)\ |\_\ C_E(A, O, R^c, R^i, R^o) \tag{1}$$

where C_E is specified by the *intension* (attributes *A*) and *extension* (objects *O*) of the concept *E* as well as its internal/input/output relations (R^c, R^i, R^o), respectively.

The *to-be* semantics in semantic algebra infers the meaning of an equivalent relation between an unknown and a known entity or concept. However, the semantics of a behavior is a *to-do* semantics embodying the meaning of an action of a person/ entity as a behavioral process which may be formally modeled by Real-Time Process Algebra (RTPA) [20].

Definition 2. The *semantics of a behavior B*, $\Theta(B)$, is a *formal process* P_B with respect to an action *A* embodied by a verb or verb phrase:

$$\Theta(B) \triangleq \Theta(J \xrightarrow{P_B} O)\ |\!>\ P_B(J, O, A, S, T) \tag{2}$$

where P_B denotes the action process *A* of the verb as well as its subject (*J*), object (*O*), space (*S*), and time (*T*).

Definition 2 leads to a formal model of the universe of discourse of human behaviors in language expression as shown in Figure 1.

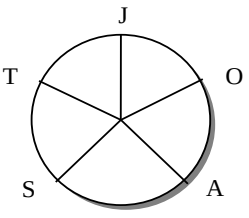


Fig. 1. The semantic space of behavioral processes for news modeling

Definition 3. The *Semantic Model of Language Expression* (SMLE) is a 5D semantic space of a behavioral process where *J* denotes the subject of the process, *O* the object, *A* the action, *S* the space, and *T* the time point when the action occurs in the semantic space.

2.2 Formal Properties of Fake News

On the basis of Definitions 1-3, the formal model of fake news or biased reports may be rigorously derived as follows.

Definition 4. A *fake news FN*|S is an expression of an entity and/or behavior that contradict to the facts in one or more dimension(s) of the 5D SMLE model:

$$\begin{aligned} FN|\mathrm{S} &\triangleq Claim|\mathrm{S} \ \dot{:}\ Fact|\mathrm{S} \neq \varnothing \\ &\triangleq Claim|\mathrm{SM}.(JP|\Xi \cup VP|\Xi \cup OP|\Xi \cup PC|\Xi \cup TC|\Xi \cup CC|\Xi \cup XC|\Xi) \\ &\quad \neq Fact|\mathrm{SM}.(JP|\Xi \cup VP|\Xi \cup OP|\Xi \cup PC|\Xi \cup TC|\Xi \cup CC|\Xi \cup XC|\Xi) \end{aligned} \tag{3}$$

where ($JP|\Xi$, $VP|\Xi$, $OP|\Xi$, $PC|\Xi$, $TC|\Xi$, $CC|\Xi$, $XC|\Xi$) denote the sets ($|\Xi$) of Part of Speeches (POS) [13] known as the subject/verb/object phrases and place/time/complement/ auxiliary clauses, respectively.

The SMLE model may be formally refined as a Structural Model (SM) of the Semantic Structures of Expressions (SSE) in Figure 2 for describing the attributes and properties of fact statements and a fake news against them. The *SSE*|SM model of the semantics of sentences formally defines the attributes of claims and facts including *Claim*|SM, *Fact*|SM, *File*|SM, *POS*|SM, *SynonymKB*|SM, *NegativeModifierKB*|SM, *DependentTree*|SM, and *Sim*|SM.

III. THE ALGORITHM OF DIFFERENTIAL SENTENCE SEMANTIC ANALYSIS

Based on the semantic theory of natural language expressions and the mathematical model of fake news as developed in the preceding section, an Algorithm of Differential Sentence Semantic Analyses (DSSA) is designed for detecting fake news against known facts by semantic analyses at sentence, paragraph, and essay levels. The algorithm of *DSSA*|PM is carried out by three processes as shown in Figure 2, i.e.: a) Claims and facts parsing; b) Differential sematic analysis; and c) Fake news determination. The DSSA algorithm is implemented by double-layer iterations where the outer layer processes each given claim of suspect news labelled by an *ID*|N (|N is a type suffix of natural number) in a set of required analyses. The inner layer analyzes each claim as well as known facts in a given dataset through processes (a) to (c) of *DSSA*|PM as explained in the following subsections. A knowledge base consists of synonyms *SynKB*|SM and negative-modifiers or antonyms *NegKB*|SM is established prior to run the *DSSA*|PM algorithm.

3.1 Claims and Facts Parsing

A suspect claim sentence as a fake news identified in the *DSSA*|PM algorithm is represented by *Claim*|S in the string type (|S); while the associated set of fact statements $\underset{k=1}{\overset{F|N}{R}} File(i)|S.Fact(k)|S$ is denoted by $F|N$ facts in the given files or searching results where the *big-R* calculus representing a recursive structure or function in RTPA [20]. In process (a), the claim and fact parsing process of *DSSA*|PM adopts a two-round sentence parses by a natural language parser, *Parser*|PM, that adopts the Stanford Parser [26] or Spacy [21]. The natural language parser scans both claims and facts in two-rounds. The claim parsing results in the partitions of each of the 7 POS phases/clauses $\underset{i=1}{\overset{n|N}{R}}\underset{k=1}{\overset{F_i|N}{R}}\underset{j=1}{\overset{7}{R}} POS(i,k,j)|\Xi$ are retained in individual sets (|Ξ). The facts parsing does the same for each of the fact sentences in $\underset{k=1}{\overset{F|N}{R}} File(i)|S.Fact(k)|S$.

```
// Algorithm of Differential Sentence Semantic Analysis (DSSA)
DSSA|PM(<I:: R_{i=1}^{C|N} Claim(i)|S>; <O:: R_{i=1}^{C|N} Sim(i)|N ∈ {(0, Negative), (1, Positive), (2, Partial)}>;
        <H:: R_{i=1}^{n|N} R_{k=1}^{F|N} SSE(i, k)|SM, R_{i 1 k 1 j 1}^{C|N F_i|N 7} POS(i, k, j)|Ξ, R_{i 1 k 1}^{C|N F_i|N} File(i)|S.Fact(k)|S,
        SynKB|SM, NegKB|SM>)    // POS - Parts of speaches
{R_{i=1}^{C|N} (R_{k=1}^{F_i|N} (→ F_i|N :— |File(i)|S.Fact(k)|S|
    // a) Parse a claim and related facts
    → Claim|L :— T|L
    ⟞ Parser|PM(<I:: i|N, Claim|L, F_i|N = 1, R_{k=1}^{F_i|N} Sentence(i,k)|S>; <O:: R_{i=1}^{n|N} R_{k=1}^{F_i|N} R_{j=1}^{7} POS(i,k,j)|Ξ >;
        <H:: R_{i 1 k 1}^{n|N F|N} POS(i, k)|SM, R_{i 1 k 1}^{C|N F_i|N} File(i)|S.Fact(k)|S>)
    → Claim|L :— F|L
    → R_{k=1}^{F|N} Sentence(i,k)|S := R_{k=1}^{F|N} File(i)|S.Fact(k)|S
    ⟞ Parser|PM(<I:: i|N, Claim|L, F_i|N, R_{k=1}^{F_i|N} Sentence(i,k)|S>; <O:: R_{i=1}^{n|N} R_{k=1}^{F_i|N} R_{j=1}^{7} POS(i,k,j)|Ξ >;
        <H:: R_{i 1 k 1}^{n|N F|N} POS(i, k)|SM, R_{i 1 k 1}^{C|N F|N} File(i)|S.Fact(k)|S>)
    // b) Differential sematic analyses (DSA)
    ⟞ DSA|PM(<I:: i|N, F_i|N, POS(i, k, j)|Ξ>, <O:: R_{k=1}^{F_i|N} Match(i,k)|N>;
        <H:: R_{i=1}^{C|N} R_{k=1}^{F|N} Diff (i, k)|SM. R_{j=1}^{7} POS(i, k, j)|Ξ; R_{k=1}^{F_i|N} File(k)|S>)
  )
  // c) Fake news determination
  → ( ❖ 1/F_i|N Σ_{k=1}^{F_i|N} Match(i,k)|N = 0
      → Sim(i)|N := 0                    // Fake news
    | ❖ 1/F_i|N Σ_{k=1}^{F_i|N} Match(i,k)|N = 1
      → Sim(i)|N := 1                    // True news
    | ~
      → Sim(i)|N := 2                    // Suspect news (partially true)
  )
}
```

Fig. 2 The Algorithm of Differential Sentence Semantic Analyses (DSSA)

3.2 Differential Analysis of Sentence Semantics

The second process of the *DSSA*|PM algorithm invokes a process of *Differential Semantic Analysis* (*DSA*|PM). *DSA*|PM carries out recursive paragraph learning among related fact files $\overset{n|\mathrm{N}}{\underset{i=1}{R}}\overset{F_i|\mathrm{N}}{\underset{k=1}{R}}\overset{7}{\underset{j=1}{R}} POS(i,k,j)|\Xi$ against a certain claim based on the parsed POS. The outcomes of *DSA*|PM are a set of differential matching scores $\overset{F|\mathrm{N}}{\underset{k=1}{R}} Match(i,k)|\mathrm{N}$ for each *i*th claim against all $F_i|\mathrm{N}$ facts. The kernel of the *DSA*|PM algorithm iteratively determines the differential scores according to Definition 5.

Definition 5. The *differential match score* $Diff|\Xi$ of fake news $Claim|\mathrm{S}$ is determined by its level of consistency with respect to known facts $Fact|\mathrm{S}$:

$$Diff|\Xi \triangleq Claim|\mathrm{S} \ominus Fact|\mathrm{S} = \frac{|Claim|\Xi \cap Fact|\Xi|}{|Claim|\Xi \cup Fact|\Xi|} \quad (4)$$

A positive semantic differentiation is determined by *DSA*|PM iff $Diff|\Xi = 0$ across all facts learnt which indicates that every semantic item between a pair of claim and fact sentences have been fulfilled. The semantic determination for a fake news in the *DSSD*|PM algorithm is aggregated from the analyses of their syntactic matches, semantic consistency, and supplemented by a macro statistical score learnt at the levels of concepts, sentences, and paragraphs from the bottom up.

3.3 Fake News Determination between Claims and Facts

In Process (c), the *DSSA*|PM algorithm determines if a given claim is a fake news in three categories: fake news (0), true news (1), and partially suspect news (2) based on the analytic score in $Diff|\Xi$ as obtained in Process (b).

The final determination is $Sim|\mathrm{N} = 0\ (\text{fake news})$ if any fact checking in $\overset{F_i|\mathrm{N}}{\underset{k=1}{R}} Match(i,k)|\mathrm{N}$ is negative. However, $Sim|\mathrm{N} = 1\ (\text{true news})$ if all $\overset{F_i|\mathrm{N}}{\underset{k=1}{R}} Match(i,k)|\mathrm{N} = 1$. Otherwise, $Sim|\mathrm{N} = 2\ (\text{patially true news})$. In practice, a partially true news is treated as a fake news in a more restricted classification. Alternatively, a threshold may be introduced to convert the partial category to separates sets of true or false news.

It is noteworthy that the negative semantic of fake mews may be detected by either its syntactic or semantic negations. The former is characterized by one or more of its modifiers of noun phrases, verb phrases, and complement phrases is/are negative or contradiction in the parsed part of speeches $POS(i,k,j)|\Xi$. However, the latter is determined based on the supporting knowledge bases of antonyms *SynKB*|SM and the negative semantics of modifier clauses *NegKB*|SM.

IV. EXPERIMENTAL RESULTS BASED ON THE DSSA ALGORITHM

Contrary to a data-driven neural network technology that may take days to train and execute testing cases, the advantage of the DSSA|PM algorithm is a unique autonomous system for training-free machine learning developed in our lab. The DSSA algorithm with supporting sub-algorithms and associated knowledge base are implemented in MATLAB. A set of 876 sample claims have been randomly selected from the DataCup'19 database [2] to evaluate the autonomous fake news detector in three categories of positive, negative, and partially suspected claims against a huge Fact Knowledge Base (FKB) provided by DataCup'19.

4.1 The System Environment of DataCup'19

The DataCup'19 competitions were established on Amazon Elastic Compute Cloud (EC2) [2]. Participants are represented by a virtual machine instances which is configured by a basic EC2 environment, a Python 3 compiler, a natural language toolkit (NLTK) [22], and a syntax parser Spacy [21]. The integrated platform is built as a software Docker [23] for results delivery to the organizer with remote communicating tools between the EC2 and the DataCup'19 organizer. EC2 is built on Ubuntu 16.04 as the operating system. This integrated Linux system includes a Python 3 compiler. Amazon provides basic virtual machine configurations for remote access. After an EC2 instance is launched, we use Windows Secure Copy Protocol (WinSCP) [24] and Putty [25] to communicate between local and remote machines. The DSSA system's performance is 20 sample analyses per minute constrained by the speed of the remote EC2 platform.

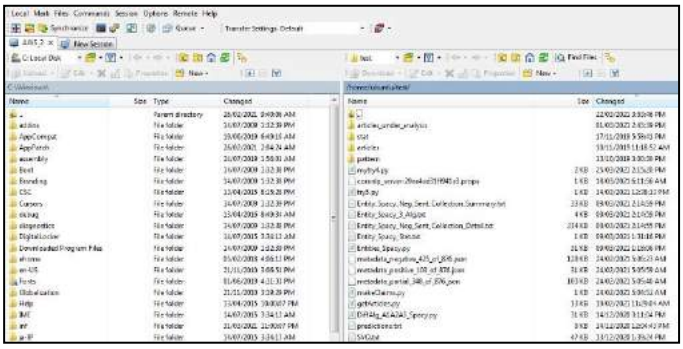


Fig. 3. The configurations of WinSCP for DataCup'19 experiments

A screenshot of the structure of WinSCP [24] is provided in Figure 3 which illustrates the scale and complexity of the integrated virtual machine. WinSCP provides a GUI-based user interface for file retrieving, manipulation, and transmission. Structures of the local and the remote folders are listed on the left/right-hand sides, respectively, by default. WinSCP includes the executable algorithm files, metadata files, and an article database of facts. Files in both locations can be easily transferred by drag-and-drop across the windows. Putty [25], on the other hand, is a command line-based interface for installing Python libraries, execute program code, and wrap program in a

docker image [23]. Special file operations such as file compression and extraction are also supported by Putty.

4.2 Experiments and Results

A sample of the testing cases in DataCup'19 is shown in Figure 4 where both the claim and facts are provided. Claim 3 is one of the challenging tests to traditional neural-network-based technologies for fake news recognition, because there are a number of minor contradictions between the claim and facts. For instances, the object "*white*" identified in the claim was not mentioned anywhere in the facts; and a few modifier clauses are different. The DSSA|PM algorithm has correctly classified Claim 3 as partially true ($Sim|N = 2$).

Claim 3: "Appalachia is the poorest country in the U.S. and happens to be more than 90 percent white."

Fact 3.1: "Appalachia had its poverty rate, 31 percent in 1960 and was 16.7 percent over the 2012–2016 period."

Fact 3.2: "Despite progress, Appalachia still does not enjoy the same economic vitality as the rest of the nation."

Fact 3.3: "Central Appalachia in particular still battles economic distress, with concentrated areas of high poverty."

Fig. 4. Sample Claim 3 against known facts in DataCup'19

The training-free DSSA algorithm is tested on a large set of 876 randomly selected samples from the DataCup'19 datasets where each of them is autonomous analyzed against several-dozen pages of fact documents. The large-scale experimental results in three categories, i.e., positive, negative, and partially suspect news claims, are reported in Figure 5, which demonstrates an overall accuracy of 70.4%. The results of accuracy are encouragingly higher than those of the top results (DT maximum) of DataCup'19 at the accuracy level about 55.0% [2] which adopted classic data-driven machine learning approaches without semantic learning and comprehension.

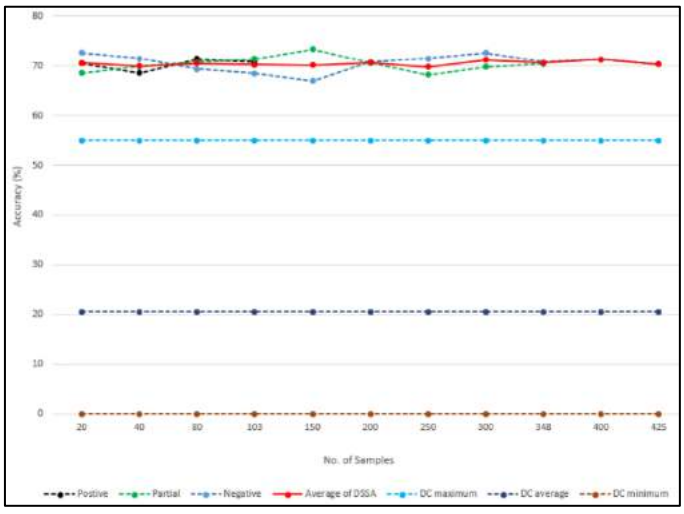


Fig. 5. Testing results generated by the DSSA algorithm

The experimental results shown in Figure 5 demonstrate that the DSSA|PM algorithm and underpinning theories enable sentence semantic differentiation in natural language to be autonomously identified and assessed. By comparative analyses, the DSSA|PM methodology has recently over performed the state-of-the-arts for fake news detection at the accuracy levels 70.4% vs. 55.0% as obtained in DataCup'19 [2]. In the upper part of Figure 5, DSSA|PM has performed better across all three categories of claims with a 15% margin. However, as shown in the lower part of Figure 5, the top three participants of DataCup'19 were ranked as 55%, 53%, and 51% in accuracy, respectively. The average rate of the top 100 teams is 20.5% because the majority were below 30% in accuracy. Among the entire 500+ participants, the average scores would be lower than 10% in DataCup'19. Therefore, the performance of the DSSA|PM-based methodology indicates a well expected advantage for training-free semantic learning and analysis for fake news recognition.

V. CONCLUSION

This work has developed an autonomous system for fake news recognition based on a novel methodology of machine learning for semantic comprehension. It has been recognized that fake news is a contradictory or partially false statement against known facts and logical consistency in semantics. It has been recognized that the field of fake news detection is dependent on a new type of machine learning known as semantic knowledge learning which is beyond traditional data-driven machine learning capability constrained by statistical regressions. The training-free machine learning algorithm of differential sematic analyses has been designed and implemented. A comprehensive set of 876 experiments randomly selected from DataCup'19 has demonstrated a 70.4% accuracy that outperform the traditional data-driven neural network technologies at the current levels ranged from 1.0% to 55.0%. The DSSA methodology has paved a way towards autonomous, training-free, non-data-driven, and real-time trustworthy technologies for machine knowledge learning and semantic manipulations.

ACKNOWLEDGEMENT

This work is supported in part by the IDEaS program sponsored by DND, Canada. The authors would like to thank the anonymous reviewers for their valuable suggestions and comments on this paper.

REFERENCES

[1] J.Z. Pan, S. Pavlova, C. Li, N. Li, Y. Li, and J. Liu (2018). Content based fake news detection using knowledge graphs. *International semantic web conference*, Springer, 669-683.

[2] DataCup (2019). https://www.datacup.ca/.

[3] A. Gupta, P. Kumaraguru, C. Castillo and C. Meier, C. (2014). Tweetcred: Real-time credibility assessment of content on twitter. *International Conference on Social Informatics*. Springer, 228–243.

[4] J. Ma, W. Gao, P. Mitra, S. Kwon, B. Jansen K. Wong, and M. Cha (2016). Detecting rumors from microblogs

with recurrent neural networks. *Proceedings of IJCAI*. 3818-3824

[5] B. Markines, C. Cattuto, and F. CMenczer (2009). Social spam detection. *Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web*. ACM, 41–48.

[6] V. Rubin, Y. Chen and N. Conroy (2015). Deception detection for news: three types of fakes. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–4

[7] K. Shu, A., Sliva, S. Wang, J. Tang, and H. Liu (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19 (2017), 22-36.

[8] Y. Wang (2017), Keynote: Cognitive Foundations of Knowledge Science and Deep Knowledge Learning by Cognitive Robots, *16th IEEE Int'l Conf. Cognitive Informatics & Cognitive Computing* (ICCI*CC 2017), Univ. of Oxford, UK, IEEE CS Press, July, pp. 4.

[9] H.A. Simon (1983), Why Should Machines Learn? in R.S. Michalskl et al. eds, *Machine Learning, an Artificial approach*, Tioga Publishing Co., 1983, pp. 25-35.

[10] Y. LeCun, Y. Bengio and G.E. Hinton (2015), Deep Learning, *Nature*, 521(7553):436-444.

[11] M. Mohri, A. Rostamizadeh, and A. Talwalkar (2012). *Foundations of Machine Learning. MIT Press, MA, USA.*

[12] W. Ferreira and A. Vlachos (2016). Emergent: a novel data-set for stance classification. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. ACL. 1163-1168.

[13] Y. Wang (2009). A formal syntax of natural languages and the deductive grammar. *Fundamenta Informaticae*. 90(4), 353–368.

[14] Y. Wang (2013). On Semantic Algebra: A Denotational Mathematics for Cognitive Linguistics, Machine Learning, and Cognitive Computing, *Journal of Advanced Mathematics and Applications*, 2(2), 145–161.

[15] E. Keenan ed. (1975). *Formal Semantics of Natural Language*, Cambridge University Press, UK.

[16] N. Chomsky (1965). *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

[17] H. Eugene ed. (1996), Cognitive linguistics in the Redwoods: The Expansion of a New Paradigm in Linguistics, *Cognitive Linguistics Research*, 6. Berlin.

[18] R.A. Wilson and C.K. Frank eds. (2001), *The MIT Encyclopedia of the Cognitive Sciences*, MIT Press, MA.

[19] Y. Wang (2015). Concept Algebra: A Denotational Mathematics for Formal Knowledge Representation and Cognitive Robot Learning, *Journal of Advanced Mathematics and Applications*, 4(1), 62–87.

[20] Y. Wang (2008). RTPA: A Denotational Mathematics for Manipulating Intelligent and Computational Behaviors. *Int'l Journal of Cognitive Informatics and Natural Intelligence*, 2(2), 44-62.

[21] Spacy project (2021), Industrial-Strength Natural Language Processing in Python, https://spacy.io/.

[22] NLTK project. (2020), Natural Language Toolkit, https://www.nltk.org/install.html.

[23] Docker (2021). https://www.docker.com/.

[24] WinSCP (2021), WinSCP, https://winscp.net/eng/index.php.

[25] Putty (2021). https://www.putty.org/.

[26] Stanford Univ. (2019), The Stanford Natural Language Processing Group, Stanford Parser, http://nlp.stanford.edu:8080/parser/index.jsp.

PROGRESS ON A PERIMETER SURVEILLANCE PROBLEM

Jeremy Avigad

Carnegie Mellon University
Department of Philosophy

Floris van Doorn

University of Pittsburgh
Department of Mathematics

ABSTRACT

We consider a perimeter surveillance problem introduced by Kingston, Beard, and Holt in 2008 and studied by Davis, Humphrey, and Kingston in 2019. In this problem, n drones surveil a finite interval, moving at uniform speed and exchanging information only when they meet another drone. Kingston et al. described a particular online algorithm for coordinating their behavior and asked for an upper bound on how long it can take before the drones are fully synchronized. They divided the algorithm's behavior into two phases and presented upper bounds on the length of each phase based on conjectured worst-case configurations. Davis et al. presented counterexamples to the conjecture for phase 1. We present sharp upper bounds on phase 2 which show that in this case the conjectured worst case is correct, and we report new lower bounds on phase 1.

Index Terms— multi-agent systems, decentralized algorithms, perimeter surveillance, small unmanned aerial vehicles (UAVs)

1. INTRODUCTION

In 2008, Kingston, Beard, and Holt [1] considered a problem in decentralized control in which a group of small unmanned aerial vehicles (UAVs) or drones is required to surveil a linear interval with changing borders. In their model, the drones all move along the interval at the same uniform speed and can exchange information only when they meet. Because the borders of the interval and the number of operant drones can change over time, the drones have imperfect information as to the global state.

Kingston et al. described an algorithm for coordinating the drones and considered the problem of bounding the time to synchronization in a setting where the parameters remain fixed. They divided the algorithm's behavior into two phases which we will call *phase 1* and *phase 2*. Normalizing units so that a single drone can traverse the interval in one unit of time, they claimed an upper bound of 3 units of time for phase 1 and an upper bound of 2 units of time for phase 2. In each case, the bounds were based on the behavior of what they took to be the worst-case starting configurations.

In 2019, Davis, Humphrey, and Kingston [2] pointed out that the previous work did not justify the claimed characterizations of the worst-case behavior, and, in fact, they provided a counterexample to the bound for phase 1. As a result, there is currently no rigorously established bound on the time to synchronization that is independent of the number of drones. In this paper, we establish sharp upper bounds on the length of phase 2, showing that the originally claimed worst-case behavior is correct. We report on improved lower bounds on the length of phase 1 as well as the total time to synchronization.

It is by now well understood that decentralized coordination of UAVs raises interesting combinatorial challenges [3, 4, 5]. What the Kingston–Beard–Holt example shows is that difficult combinatorial problems arise even when dealing with fairly simple models, and that new mathematical ideas and techniques are needed to handle them. Our results here are intended as a contribution to the mathematical toolbox.[1]

2. THE PROBLEM

To describe the model under consideration, it is convenient to normalize units so that the drones are surveilling the unit interval $[0, 1]$ and moving at a velocity of one unit per unit time. At each point in time, each drone has a direction $d = \pm 1$, where 1 indicates that the drone is moving to the right and -1 indicates that it is moving to the left. Each drone also has an estimate of the form $((a, \ell), (b, m))$ where a is the left endpoint of the interval, ℓ is the number of drones to the left, b is the right endpoint, and m is number of drones to the right. Kingston, Beard, and Holt [1] wanted to consider a scenario where the data keeps changing, so these estimates may be wrong; in particular, a and b do not need to be in the unit interval. Each drone recognizes the leftmost or rightmost border when it reaches it. Two drones can only exchange information when they meet, that is, occupy the same position in the interval.

Work supported in part by AFOSR grant FA9550-18-1-0120 and Sloan Foundation grant G-2018-10067

Work supported in part by Sloan Foundation grant G-2018-10067

[1] David Greve at Collins Aerospace has independently obtained a substantially different proof of our Theorem 1 with 2 replacing $2 - 1/n$ [6], and he has recently formalized the proof presented here (personal communication) in the ACL2 verification system [7].

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Suppose we have n drones on the unit interval, numbered from left to right $1, \ldots, n$. The ith drone's *left endpoint* is $(i-1)/n$, its *right endpoint* is i/n, and its *interval* is the closed interval with those endpoints. We say the *common endpoint* of drones i and $i+1$ is the right endpoint of drone i, which is equal to the left endpoint of drone $i+1$. The desired situation is that each drone remains in its interval, moving back and forth between its left and right endpoints.

Kingston et al. proposed the following algorithm to attain this behavior. Write $(\alpha, \beta) = ((a, \ell), (b, m))$ for the drone's estimates. Based on this data, each drone can calculate the interval it *thinks* it is supposed to surveil as follows:

- The size of the interval is $I(\alpha, \beta) = (b-a)/(\ell+m+1)$, that is, the length of the estimated interval divided by the number of drones.
- The left endpoint is

$$L(\alpha, \beta) = a + \ell I(\alpha, \beta) = b - (m+1)I(\alpha, \beta)$$

- The right endpoint is

$$R(\alpha, \beta) = a + (m+1)I(\alpha, \beta) = b - nI(\alpha, \beta)$$

According to the algorithm, each drone continues in the direction it is moving until one of these events occurs:

- If a drone hits the left border, it updates its left estimate with the correct left endpoint of the interval and the fact that there are no drones to the left, and then it turns around. The case where a drone hits the right border is handled similarly.
- If two drones meet, the left one adopts the right estimate from the drone to its right (adding 1 to the number of drones to the right), and vice-versa. As a result, the two drones agree as to their estimates of the intervals they are supposed to surveil. The two then set their directions so that they are headed to their common endpoint.
- If two drones are traveling together (with consistent estimates) and reach their common endpoint, they split, i.e. one of them reverses direction in order to stay in its estimated interval.

We assume that when two or more drones start together, they all share their estimates at time 0. We call the first type of event a *border* event, the second type of event a *meet* event, and the third type of event a *separation* event. Notice that, as a special case, the second and third can happen simultaneously, if two drones meet at their common endpoint. We call that a *bounce* event. Fig. 1 depicts a sample run of the algorithm with five drones, with the interval extending left to right and time flowing downward.

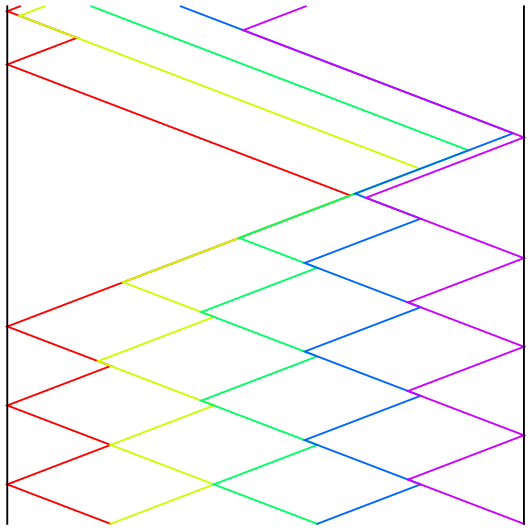


Fig. 1. A sample run of the algorithm with $n = 5$ drones.

The arguments below can be made rigorous by formalizing the notion of a *configuration* (that is, a time, t, and the sequence of positions, directions and estimates of all the drones at that time), and, given a configuration, the next configuration at which one of the events above occurs. It then makes sense to talk about the sequence of eventful configurations from a given start configuration, and all the informal claims below can be interpreted in terms of that.

Given a particular start configuration, say a drone is *left synchronized* at time t if beyond that point it never goes to the left of its left endpoint, and similarly for *right synchronized*. A drone is *synchronized* at time t if it is left and right synchronized. Kingston et al. conjectured the following:

Conjecture 1. *From any start configuration, all drones have correct estimates by time 3.*

Conjecture 2. *If all drones have correct estimates at time t, then all drones are synchronized by time $t+2$.*

We call the time between the start and the moment that all drones have correct estimates *phase 1*, and the time after phase 1 until the moment that the drones are synchronized *phase 2*.

Kingston et al. sketched a proof of each conjecture, in each case based on a claims that a certain start configurations gave rise to the worst-case behavior. The two conjectures imply that for any start configuration, the drones are synchronized by time 5.

Davis, Humphrey, and Kingston [2] showed that Conjecture 1 is false, by exhibiting counterexamples with $n = 3$ that require up to $3 + 1/2$ units of time before all the drones have correct estimates. For that purpose, they used the AGREE model checker [8], which required fixed bounds on all the parameters. In particular, they had to limit the estimates of the

number of drones to the left or right at 20. With those restrictions, the tool reported upper bounds of $3 + 2/3$ on the time until all three drones have complete information, and $4 + 1/3$ units of time until full synchronization. The tool also reported absolute upper bounds of 2 on phase 2, with $n \leq 6$; they report that the verification for $n = 6$ required about 20 days of computation using 40 cores. They do not report any results for larger n. In particular, there was no rigorously established bound on the length of either phase, or total time to synchronization, that is independent of n.

Conjecture 2 is clearly implied by the following statement: if all drones start with correct estimates then they are synchronized by time 2. The implication follows, because we can consider the configuration at time t as the new start configuration. In Section 3, we prove the following:

Theorem 1. *Assuming all the drones have the correct estimates, they are all synchronized at time* $2 - 1/n$.

This shows that the conjecture by Kingston et al. as to the worst-case configurations is correct.

In an extended version of this paper,[2] we obtain the following additional information:

Theorem 2. *If all drones start with incorrect estimates, and they all have correct estimates at time* t*, then all drones are synchronized by time* $t + 1 - 1/n$.

We also improve the lower bounds as follows:

Theorem 3. *For every* $n \geq 3$ *and* $\varepsilon > 0$*, there is a start configuration such that drones do not have correct estimates before time* $4 - 1/n - \varepsilon$*, and are not fully synchronized before time* $5 - 3/n - \varepsilon$.

This strengthens the counterexample obtained by Davis et al. In the extended version we also describe progress toward an upper bound for phase 1.

We mention in passing that the case $n = 1$ is trivial; a single drone is already synchronized, though it may not have correct estimates until time 2. The case $n = 2$ is also easy to analyze; drones have correct estimates by time 2 and are synchronized by time $2 + 1/2$. Both these bounds are sharp, so $n = 3$ is the first interesting case.

It is important to note that Kingston et al. were not looking for an algorithm to synchronize all the drones as quickly as possible. Rather, they were independently interested in the behavior of that particular algorithm for updating information in the face of changing borders and addition or subtraction of drones. Given that, the question about worst-case behavior even under fixed conditions is natural.

3. AN UPPER BOUND ON PHASE 2

Once the drones have the correct estimates as to the left and right endpoints and the number of drones on either side, each drone knows its proper interval, and the behavior of the algorithm from that point on can be described more simply: when two drones meet, they escort each other to their common endpoint and then separate. Our goal is to prove Theorem 1, which guarantees that the drones are all synchronized within $2 - 1/n$ units of time.

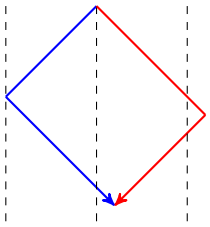


Fig. 2. A picture proof of Lemma 2. The two intervals between the vertical lines indicate the intervals of drone j (blue) and drone $j + 1$ (red).

By symmetry, it suffices to show that all the drones are left synchronized by time $2 - 1/n$. Kingston, Beard, and Holt [1] gave a short argument that all drones are eventually left synchronized, although the bound that is implicit in that argument is linear in n. The argument goes as follows: suppose at some time, t, drones $1, \ldots, j$ are left synchronized. Eventually, drone j will meet drone $j + 1$, and then they will travel to their common endpoint and separate. It suffices to show that at this point, $j + 1$ in left synchronized, because then by induction we have that all drones are eventually left synchronized. We present their proof of this in Lemma 2. Our proof of Theorem 1 is based on a subtle refinement of their argument.

Lemma 1. *Suppose that at time* t *drone* j *is moving to the right. Then the next time drone* j *changes direction, it is at or to the right of its right endpoint. The same is true with "right" replaced by "left."*

Proof. If drone $j < n$ is moving to the right, the only two events in which it can change direction is when meeting drone $j + 1$ or separating from drone $j + 1$. If the next time drone j turns left is when meeting drone $j + 1$, then at that point they are to the right of their common endpoint, which is the right endpoint for drone j. If the next time drone j turns left is when separating from (or bouncing off) drone $j + 1$, then at that point they must both be at their common endpoint. If drone n is moving to the right, it only changes direction at a border event, which is at its right endpoint. The last observation follows by symmetry. □

Lemma 2. *Suppose at time* t*, drone* j *is left synchronized and drone* j *and* $j + 1$ *separate at their common endpoint. Then drone* $j + 1$ *is left synchronized at time* t.

[2] https://arxiv.org/abs/2008.04262

Proof. By induction, we show that every subsequent meet, bounce, and separation event involving j and $j+1$ occurs to the right of their common endpoint.

After the separation, drone j is moving to the left. By Lemma 1, the next time it changes direction, it is at or to the left of its left endpoint. Since it is left synchronized, we know it is *at* its left endpoint. Similarly, after the separation, drone $j+1$ is moving to the right, and the next time it changes direction is it at or to the right of its right endpoint. So the next time drone j and $j+1$ meet, they are at or to the right of their common endpoint, since drone $j+1$ must has taken at least as long to turn around as drone j; see Fig. 2 for a visual depiction. They then travel left to their common endpoint and separate, and the situation repeats. □

We now draw out two useful consequences of Lemma 2:

Lemma 3. *Suppose at time t drone j and $j+1$ are together moving to the left, and drone j is left synchronized. Then drone $j+1$ is also left synchronized at time t.*

Proof. If they are together and moving to the left, they are to the right of their common endpoint. Eventually they will reach their common endpoint and separate, say, at time t'. By Lemma 2, drone $j+1$ is left synchronized at time t'. But since drone $j+1$ is to the right of its left endpoint between time t and t', it is in fact left synchronized at time t. □

Lemma 4. *Suppose at time t drone $j < n$ is at or to the right of its right endpoint, moving right, and left synchronized. Then drone $j+1$ is left synchronized at time t.*

Proof. Since drone j is moving right and is to the right to its right endpoint, it cannot be together with drone $j+1$. Eventually drone j and $j+1$ will meet to the right of their common endpoint, say at time t', and then they will move left together. At that point, by Lemma 3, drone $j+1$ is left synchronized. Since drone $j+1$ is to the right of its left endpoint between time t and t', it is already left synchronized at time t. □

We have now arrived at the key refinement of the argument by Kingston et al. We say that drones j and $j+1$ *have met* by time t if either they started together, moving in the same direction, or they have been involved in a meet or bounce event. It turns out that we have much more information about the behavior of the drones once this is the case. Fortunately, it is not hard to show that this happens within one unit of time, for all the drones uniformly.

Lemma 5. *For every $j < n$, drones j and $j+1$ have met by time 1.*

Proof. Intuitively, the worst case is where j and $j+1$ start close together with j moving to the left and $j+1$ moving to the right. Eventually, j turns around at or before it reaches 0 and $j+1$ turns around at or before it reaches 1, and then j and $j+1$ will meet. At that point, together they have traveled at most the twice the length of the interval, which means that each one has traveled at most one unit of distance.

We can make this argument more rigorous as follows. Suppose drone j starts at position x and drone $j+1$ starts at position $y \geq x$. Furthermore, let w be the position of drone j when it first moves right (so $w = x$ if j starts moving right, and otherwise w is the position where drone j first turns around), and let z be the position of drone $j+1$ when it first moves left. Then the total distance traveled by both drones before they meet is $2(z-w)-(y-x)$, which means the drones meet at time $z - w - (y-x)/2 \leq z - w \leq 1$. □

Lemma 6. *Suppose that at time t, drone j is moving to the left and drone $j+1$ is not together with drone j. Suppose also that j and $j+1$ have met by time t. Let t' be the last time before time t that drones j and $j+1$ bounced or separated. Then j has been moving left since time t'.*

Proof. If at some point between t' and t drone j was moving to the right, something must have turned it to the left. But that can only have been a meet or separation or bounce event. If it was a meet event, the fact that j and $j+1$ are not together at time t means there was also a separation event. Both situations contradict the fact that t' is the last time before time t that drones j and $j+1$ bounced or separated. □

Lemma 7. *Suppose $j < n$ and at time t, drones $1, \ldots, j$ are left synchronized and drone j and $j+1$ have met. Then at time $t + 1/n$, drones $j+1$ is left synchronized as well.*

Proof. Suppose drone j is left synchronized. If it is moving to the right, it will be at or to the right of its right endpoint within time $1/n$, possibly having met drone $j+1$ along the way. At that point drone $j+1$ is left synchronized, by Lemmas 2 and 4. If drone j is moving to the left and it is together with drone $j+1$, drone $j+1$ is left synchronized at time t by Lemma 3.

Finally, suppose drone j is moving to the left and is not together with drone $j+1$. Since we are assuming drones j and $j+1$ have met by time t, there is a $t' < t$ where drones j and $j+1$ bounced or separated last. By Lemma 6, drone j has been moving left since time t'. Since drone j is at or to the right of its left endpoint at time t, it was to the right of its left endpoint between time t' and t. Since drone j is left synchronized at time t, this shows that it was already left synchronized at t'. By Lemma 2, drone $j+1$ was also left synchronized at time t', and hence is left synchronized at time t. □

Since drone 1 is always left synchronized and all the drones have met by time 1, by induction on $i < n$ we have that drones $1, \ldots, i$ are left synchronized at time $1+(i-1)/n$. Taking $i = n$ yields Theorem 1. It is not hard to show that this theorem is sharp.

4. REFERENCES

[1] Derek B. Kingston, Randal W. Beard, and Ryan S. Holt, "Decentralized perimeter surveillance using a team of UAVs," *IEEE Trans. Robotics*, vol. 24, no. 6, pp. 1394–1404, 2008.

[2] Jennifer A. Davis, Laura R. Humphrey, and Derek B. Kingston, "When human intuition fails: Using formal methods to find an error in the 'proof' of a multi-agent protocol," in *Computer Aided Verification (CAV) 2019*, Isil Dillig and Serdar Tasiran, Eds. 2019, pp. 366–375, Springer.

[3] Lubomír Bakule, "Decentralized control: An overview," *Annual Reviews in Control*, vol. 32, no. 1, pp. 87 – 98, 2008.

[4] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 376–386, 2005.

[5] P. B. Sujit and Randy Beard, "Multiple UAV exploration of an unknown region," *Ann. Math. Artif. Intell.*, vol. 52, no. 2-4, pp. 335–366, 2008.

[6] David Greve, "A hierarchical proof of DPSS-A," in progress, 2021.

[7] Robert S. Boyer and J Strother Moore, *A Computational Logic Handbook*, Academic Press international series in formal methods. Academic Press, second edition, 1998.

[8] Darren D. Cofer, Andrew Gacek, Steven P. Miller, Michael W. Whalen, Brian LaValley, and Lui Sha, "Compositional verification of architectural models," in *NASA Formal Methods (NFM) 2012*, Alwyn Goodloe and Suzette Person, Eds. 2012, pp. 126–140, Springer.

REAL-TIME LEARNING FOR THz RADAR MAPPING AND UAV CONTROL

Anna Guerra [†], *Francesco Guidi* [⋆], *Davide Dardari* [†], *Petar M. Djurić* [◇]

[†] DEI, University of Bologna, Italy. E-mail: {anna.guerra3, davide.dardari}@unibo.it
[⋆] CNR-IEIIT, National Council Research of Italy, Italy. E-mail: francesco.guidi@ieiit.cnr.it
[◇] ECE, Stony Brook University, New York. E-mail: petar.djuric@stonybrook.edu

ABSTRACT

In this paper we consider a joint detection, mapping and navigation problem by an unmanned aerial vehicle (UAV) with real-time learning capabilities. We formulate this problem as a Markov decision process (MDP), where the UAV is equipped with a THz radar capable to electronically scan the environment with high accuracy and to infer its probabilistic occupancy map. The navigation task amounts to maximizing the desired mapping accuracy and coverage and to decide whether targets (e.g., people carrying radio devices) are present or not. With the numerical results, we analyze the robustness of the considered Q-learning algorithm, and we discuss practical applications.

Index Terms— Autonomous Navigation, Reinforcement Learning, Q-learning, Unmanned Aerial Vehicles.

1. INTRODUCTION

Perception and cognition are two essential features for next generation radar systems. A cognitive radar (CR) is able to learn from the environment and to adjust its behaviour based on the received rewards or penalties that represent a feedback on the CR actions [1].

More recently, in [2, 3] a massive multiple-input multiple-output (MIMO) CR has been investigated for multi-target detection using a reinforcement learning (RL) algorithm. In these papers, no prior information about the statistical model of the disturbance, or of the number of targets, was assumed for the proper functioning of the radar. Following a similar research direction, [4] showed the optimization of the trajectory of a unmanned aerial vehicle (UAV)-radar for environment mapping and detection using a RL approach where rewards were predicted within a finite temporal horizon. Indeed, *time* is a key aspect for UAV networks because of their limited energy autonomy [5–7] and, thus, it should be properly accounted for when designing the UAV control for time-critical applications (e.g., search–and–rescue). In [5], an information-seeking algorithm is developed for extraterrestrial exploration and return-to-base application, whereas in [8, 9] a similar problem is solved using RL for source localization. Algorithms for UAVs formation, navigation and self-localization have been proposed in [10–14], and RL for enhancing communications has been studied in [15–18].

The advent of sixth generation (6G) cellular systems fosters the exploitation of new frequency bands, which suggests the importance to investigate indoor detection and mapping using Terahertz (THz) radar technologies, as they are expected to guarantee unprecedented levels of radio localization accuracy [19]. The advantage of operating at THz rather than microwaves is that the surface illuminated by the interrogation signal reflects back in different directions (*diffuse scattering*) and not just specularly [20]. Beyond 100 GHz, the diffuse scattering is comparable with the specular component, allowing to improve the reconstruction of the surrounding thanks to the richer backscattered signal.

In this paper, our aim is to explore this technology in the context of CR-UAV. To be successful in indoor detection and mapping, the CR-UAV has to autonomously decide where to go to improve the detection task within a limited available time. Increasing the ambient awareness through mapping can also accelerate the overall learning process and the completion of the UAV primary task. Thus, starting from [4], valid when empirical models are available, we consider a THz radar exploiting a Q-learning algorithm with a combination of intrinsic (mapping) and extrinsic (detection) rewards. Finally, we show the impact of the THz radar parameters on the attainable performance through a simulation analysis.

2. PROBLEM FORMULATION

The UAV trajectory is designed to maximize the target detection, mapping accuracy and coverage subject to the mission time T_{M} and collision avoidance. We formulate the optimization problem as a Markov decision process (MDP). This problem can be solved using a model-free RL method. An example of an indoor environment is shown in Fig. 1.

Markov Decision Processes: Following the same notation as in [21], a MDP is defined by a tuple containing the state space $\mathcal{S}$, the action space $\mathcal{A}$, the reward space $\mathcal{R}$, and the probability of transitioning from one state s_k, at time instant k, to the next state s_{k+1}. Notably, the random state at time instant k, indicated with S_k, represents the knowledge about the environment available to the agent at time instant k,

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

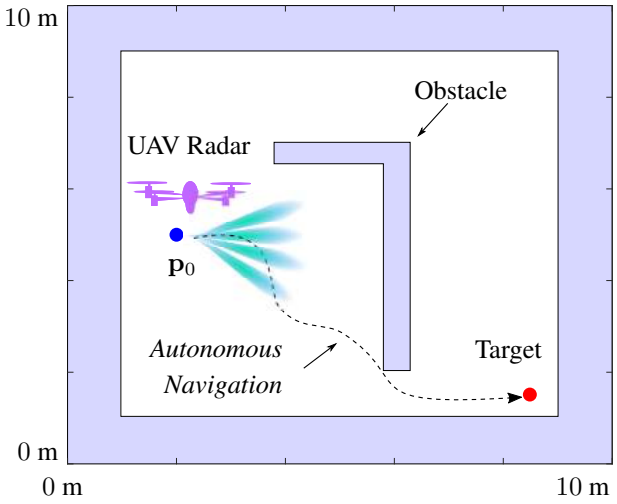


Fig. 1: Considered UAV scenario and reference map.

and it can take values $s_k \in \mathcal{S}$. The actions are chosen according to a specific policy $\pi(a_k|s_k)$, which is referred to as a probability density function (pdf) of an action.[1] The optimal policy selects actions that maximize a value function by

$$\pi^*(a_k|s_k) = \arg\max_{a_k} Q_\pi(s_k, a_k), \tag{1}$$

where the Q-function, $Q_\pi(\cdot)$, is the expected sum of discounted rewards over all possible policies and is given by

$$Q_\pi(s_k, a_k) = \mathbb{E}_\pi\left\{\sum_{l=0}^{\infty} \gamma^l R_{k+l+1} \middle| S_k = s_k, A_k = a_k\right\}, \tag{2}$$

with $0 \leq \gamma \leq 1$ being the discount rate and where the expected reward at time instant $k+1$, is $r_{k+1}(s_k, a_k) = \mathbb{E}[R_{k+1}|S_k = s_k, A_k = a_k]$. Optimal policies share the same *optimal action-value function* defined as

$$Q^*(s_k, a_k) = \arg\max_{\pi} Q_\pi(s_k, a_k), \quad \forall s_k, \forall a_k. \tag{3}$$

State: The state vector $\mathbf{s}_k$ at time k contains the UAV location, the map of the environment and a detection variable, i.e., $\mathbf{s}_k = [\mathbf{p}_k, \mathbf{m}_k, \mathrm{t}_k]^\mathsf{T}$, where $\mathbf{p}_k = [x_k, y_k]^\mathsf{T} \in \mathbb{R}^2$ is the true UAV position, $\mathbf{m}_k \in \mathbb{B}^{N_\text{cell}}$ is the true map at time k described as a vector of N_cell cells in which the map is discretized, and $\mathrm{t}_k \in \mathbb{B}$ is the target variable (equal to one if the target is present and zero otherwise). As the environment is considered stationary, it is $\mathrm{t}_k = \mathrm{t}$ and $\mathbf{m}_k = \mathbf{m}, \forall k$, with $\mathbf{m} = [m_1, \ldots, m_i, \ldots, m_{N_\text{cell}}]^\mathsf{T}$, containing the occupancy value of each cell, i.e., $m_i \in \mathbb{B}$, and N_cell being the total number of cells. The state space is[2]

$$\mathcal{S} = \underbrace{\mathbb{R}^2}_{\text{UAV position}} \times \underbrace{\mathbb{B}^{N_\text{cell}}}_{\text{Map}} \times \underbrace{\mathbb{B}}_{\text{Target}}. \tag{4}$$

[1] Note that π for a discrete state-action is a probability mass function.

[2] When the dimension of the state space is large (e.g., for large outdoors), policy iteration might suffer for the "curse of dimensionality" [22].

Actions: The UAV navigation actions can be defined as $\mathbf{a}_k = \Delta\mathbf{p}_k = [\Delta x_k, \Delta y_k]^\mathsf{T} \in \mathbb{R}^2$ in terms of position displacement $\Delta\mathbf{p}_k$ according to $N_\text{a} = 4$ actions, where the action space, for steps of Δ, is

$$\mathcal{A} = \Big\{\underbrace{[\Delta, 0]}_{\text{Right}}, \underbrace{[-\Delta, 0]}_{\text{Left}}, \underbrace{[0, \Delta]}_{\text{Up}}, \underbrace{[0, -\Delta]}_{\text{Down}}\Big\}. \tag{5}$$

Rewards: Following the *information foraging* philosophy [23, 24], we consider an extrinsic reward that is task-specific (detection) and it maps state-action pairs to a real-valued reward, and an intrinsic reward that only indirectly depends on the world state via the UAV internal belief of the state [23]. Intrinsic rewards are usually used for *reward shaping*, for example in situations with sparse rewards. The combination of intrinsic and extrinsic rewards allows to speed up the learning process and to get better policies. According to this formulation, the reward is defined as [23]

$$r_{k+1} = r_{\text{i},\,k+1} + \eta\, r_{\text{e},\,k+1}, \tag{6}$$

where we omitted the state and action dependence, η is a normalizing factor, $r_{\text{i},\,k+1} = r_{\text{c},k+1} + r_{\text{m},k+1}$ is an intrinsic reward used for obtaining a sufficient knowledge of the surrounding environment, and $r_{\text{e},\,k+1} = r_{\text{d},k+1}$ is a reward for the considered UAV task. More specifically, $r_{\text{d},k+1}$ is defined as the reward accounting for the detection rate that is

$$r_{\text{d},k+1} = \mathcal{Q}_h(\sqrt{\lambda_k}, \sqrt{\xi}), \tag{7}$$

where $\mathcal{Q}_h$ is the Marcum's $\mathcal{Q}$-function of order h, λ_k is the measured signal-to-noise ratio (SNR) at time instant k and ξ is the considered signal detection threshold [4, (37)], [25, 26].

For each radar position, we also define a mapping reward both in terms of coverage ($r_{\text{c},k+1}$) and accuracy ($r_{\text{m},k+1}$) as

$$r_{\text{c},k+1} \triangleq \frac{\sum_{i\in\mathcal{I}_k} \mathbf{1}(i \in \mathcal{D}_k)}{N_\text{cell}}, \quad r_{\text{m},k+1} \triangleq \frac{H_{k+1|k}(\mathbf{m})}{|\mathcal{I}_k|}, \tag{8}$$

where $\mathcal{D}_k \subseteq \mathcal{I}_k$ represents the subset of the indices of the intercepted cells that are discovered for the first time, and $\mathcal{I}_k$ is the set of the indices of all the cells illuminated by the radar at the kth time slot, and $\mathbf{1}(x) = 1$ if the logical condition x is verified, otherwise it is 0. Considering $r_{\text{m},k+1}$, it holds

$$H_{k+1|k}(\mathbf{m}) = -\sum_{i\in\mathcal{I}_k} b_{k+1|k}(m_i)\, \log_2\left(b_{k+1|k}(m_i)\right), \tag{9}$$

where $H_{k+1|k}(\mathbf{m})$ represents the entropy indicating the level of lack of information about $\mathbf{m}$, $|\mathcal{I}_k|$ is the cardinality of $\mathcal{I}_k$ [4, (35)], and $b_{k+1|k}(m_i)$ is the predicted belief of occupancy state of the ith cell at time slot k. Note that such reward is designed in a way to favor actions that reduce the uncertainty about the environment in the shortest possible time. Finally we consider a numerical penalty for avoiding crashes with obstacles and targets.

3. STATE ESTIMATION AND CONTROL

The CR on UAV is a system comprising two estimation processes. The first is a "*State Estimator*" that implements an occupancy grid (OG) for mapping and a detection module that determines if a target is present. The second step is a "*Policy Estimator*" for the UAV navigation.

3.1. State Estimator: Mapping with OG

The map of the environment is estimated using an OG algorithm [4], and energy measurements collected by the radar from each steering direction and different tested distances, according to the model described in [4, (13)] and [27, (35-37)].

Let $b_k(m_i)$ be the belief of the occupancy state of the ith cell at time instant k. Given the binary nature of m_i and to avoid numerical instability, the OG uses log-odds, defined as $\ell_k(m_i) \triangleq \log\left(\frac{b_k(m_i)}{1-b_k(m_i)}\right)$. The major steps are summarized as follows.

Initialization: The belief of each cell composing the map is initialized as $b_0(m_i) = 0.5$ (complete uncertainty).

Measurement Update: A new energy matrix is collected for each steering direction and time bin and it is compared with the expected received power, evaluated according to the THz scattering model of [20] and the actual knowledge of the map. More specifically, it accounts for the scattering term

$$\rho = 8\,\pi \frac{S^2\,L\,\cos(\theta_i)}{F_{\alpha_\mathrm{r}}}\left(\frac{1+\cos(\Psi)}{2}\right)^{\alpha_\mathrm{r}}, \qquad (10)$$

where S is the scattering coefficient, θ_i is the incident angle with respect to the normal of the obstacle, $\Psi = \theta_s - \theta_r$ is the difference between the reflected (θ_r) and the scattered (θ_s) angles, and L is the length of the scattering object. F_{α_r} is a scaling factor, and α_r is the width of the scattering lobe.

Hence, the likelihood functions for the case of occupied/free cells (i.e., $p(\mathbf{o}_k|m_i)$) are computed as in [4, (22-23)], where $\mathbf{o}_k$ is the observation collected at the kth instant.

Log-Odd Update: Finally, for each time instant, the log-odd update is

$$\ell_k(m_i) = \log\left(\frac{p(\mathbf{o}_k|m_i)}{1-p(\mathbf{o}_k|m_i)}\right) + \ell_{k-1}(m_i). \qquad (11)$$

3.2. Policy Estimator: Control with Q-learning

Q-learning is an off-policy temporal-difference (TD) control algorithm approach where the policy is learnt run-time while the UAV is navigating the environment. It is a model-free tabular algorithm whose main steps are reported in Alg. 1, where we included the possibility of choosing a random action with probability ϵ (ϵ-greedy approach). TD methods use a generalized policy iteration (GPI) mechanism to alternatively estimate the optimal policy in (1) and the optimal Q-value in (3).

Algorithm 1 Q-Learning Navigation for a Single Episode

Parameters: Set $(\gamma, \alpha, \epsilon)$ and the mission time T_M;
Initialization: Initialize the Q-table to zeros, and $\mathbf{s}_0$;
while $k < T_\mathsf{M}$ **do**
  Generate a random value ϵ_k;
  if $\epsilon_k < \epsilon$ **then**
    Choose a random action $\mathbf{a}_k \in \mathcal{A}$; *(exploration)*
  else
    Choose a greedy action $\mathbf{a}_k \in \mathcal{A}$ that corresponds to the maximum Q-value in $Q(\mathbf{s}_k, :)$; *(exploitation)*
  end
  UAV moves to the new state, collects the reward r_{k+1} and updates the Q-table according to (12).
end

The advantages of using TD methods instead of Monte Carlo or dynamic programming is that there is no need of a model for the environment's dynamics and an update of the return is made at each time step.

Moreover, a sample return is considered instead of the expected return in (2) by the use of sample episodes. For discrete states and actions, the Q-value in (2) can be represented by a Q-table that, at each time instant, is updated as [21]

$$Q(\mathbf{s}_k, \mathbf{a}_k) \leftarrow Q(\mathbf{s}_k, \mathbf{a}_k) + \qquad (12)$$
$$+ \alpha\left[r_{k+1} + \gamma \max_{\mathbf{a}} Q(\mathbf{s}_{k+1}, \mathbf{a}) - Q(\mathbf{s}_k, \mathbf{a}_k)\right],$$

where α is the learning rate, and the max operator is used to have a greedy policy. In this case, the learned action-value function directly approximates the optimal action-value function in (3), independently from the policy being followed.

4. CASE STUDY

We now assess the navigation and mapping performance by accounting for a realistic propagation environment and different radar parameters. For the THz scattering model, we set $S = 0.5$ (rough surface), $L = 0.5$ and $\alpha_\mathrm{r} = 1$ [20]. Then, we considered an effective radiated isotropic power (EIRP) of 30 dBm, a receiver noise figure of 4 dB, a transmitted signal with central frequency of 140 GHz, and 1 GHz bandwidth. The mapping is performed by a radar equipped with an antenna array of 100 antennas such that 10 steering directions are required for scanning the environment, and with a reading range (RR) that is alternatively set to 3 m and 7 m. The radar is initially assumed to be in $\mathbf{p}_0 = (2, 5)$ m and it moves with steps of $\Delta = 0.5$ m, equal to the cell width. For mapping parameters, we refer to [4]. For the detection module, we considered an antenna with 0 dBi gain, a RR of 7 m and a target always present and located alternatively in $(8.5, 1.5)$ m, and $(8.5, 8.5)$ m. We set ξ in (7) by considering a desired false alarm probability of 10^{-3}. We fixed $T_\mathsf{M} = 400$, $N_\mathsf{ep} = 20$ episodes, $\gamma = 0.99$, $\alpha = 0.9$, and

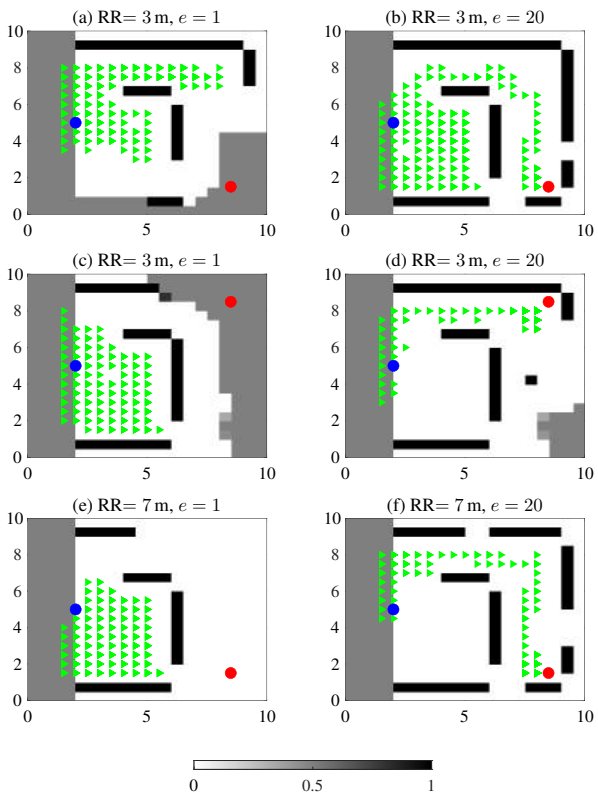


Fig. 2: Examples of estimated trajectories and maps for $e = 1$ (left) and $e = 20$ (right). Blue and red markers indicate $\mathbf{p}_0$ and the target position, respectively.

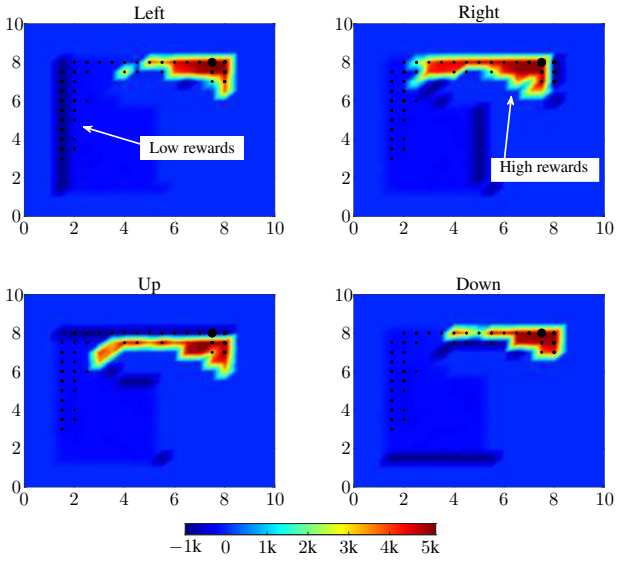


Fig. 3: Q-table related to Fig. 2-(e) with $k = T_{\mathsf{M}}$.

$\epsilon = \epsilon_{k,e}$, with $\epsilon_{k,e} = 0.2$, $\forall e > N_{\mathsf{ep}}/2$, otherwise it holds $\epsilon_{k,e} = [0.8, 0.6, 0.5, 0.3]$ for $k < [T_{\mathsf{M}}/4, T_{\mathsf{M}}/2, 3T_{\mathsf{M}}/4, T_{\mathsf{M}}]$.

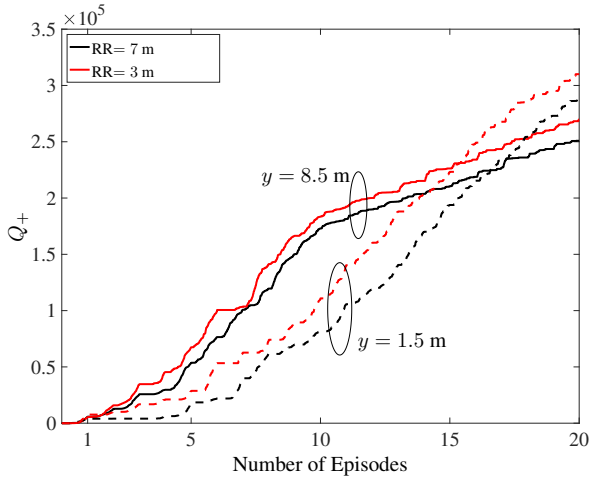


Fig. 4: Positive Q-values within a window of $N_{\mathsf{ep}} = 20$.

4.1. Results

Figure 2 shows the UAV trajectory with green markers for two different episodes, that is, $e = 1$ (left) and $e = 20$ (right), and for different radar RRs, that are RR $= 3$ m (top, middle) and RR $= 7$ m (bottom). According to the results, the UAV is capable of reconstructing a reliable copy of the map (see the reference map in Fig. 1) and of finding a good trajectory after some training episodes. In fact, during the first episode, i.e., for $e = 1$, the radar is still in an exploratory phase, as evidenced by the scarce map reconstruction, and by the followed non–optimized trajectory. This can be explained by the fact that the detection reward is sparse in the environment and mapping rewards tend faster to zero, especially for high RR. Fig. 3 reports the Q-table related to the last instant of Fig. 2-(e) for each possible action. For example, a UAV located in $(4, 7)$ will receive the highest reward by choosing the right action. By contrast, a UAV in $(2, 5)$ will receive the lowest reward by choosing the left action. Finally, Fig. 4 reports the behavior of the positive Q-values as a function of the number of episodes. Notably, for shorter RR, the UAV, driven by curiosity, is pushed to explore more, thus increasing the amount of received rewards.

5. CONCLUSION

In this paper we showed the UAV capability for autonomous navigation of an environment to accomplish the goal of detecting a target and of reconstructing a map of the indoors. We considered a Q-learning approach with a combination of intrinsic and extrinsic rewards. Our results show the possibility of attaining the objective by means of a THz radar, which augments its ambient awareness at each episode and improves its capability of accomplishing the assigned task of target detection.

6. REFERENCES

[1] S. Haykin, "Cognitive radar: a way of the future," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 30–40, 2006.

[2] A. M. Ahmed et al., "A reinforcement learning based approach for multi-target detection in massive MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, pp. 1–1, 2021.

[3] P. Liu et al., "Decentralized automotive radar spectrum allocation to avoid mutual interference using reinforcement learning," *IEEE Trans. Aerosp. Electron. Syst.*, 2020.

[4] A. Guerra et al., "Reinforcement learning for UAV autonomous navigation, mapping and target detection," in *Proc. IEEE/ION Pos. Loc. Nav. Symp.*, 2020, pp. 1004–1013.

[5] S. Zhang, R. Raulefs, and A. Dammann, "Location information driven formation control for swarm return-to-base application," in *Proc. European Signal Process. Conf.* IEEE, 2016, pp. 758–763.

[6] F. Koohifar et al., "Autonomous tracking of intermittent RF source using a UAV swarm," *IEEE Access*, vol. 6, pp. 15884–15897, 2018.

[7] Emanuel Staudinger et al., "The role of time in a robotic swarm: A joint view on communications, localization, and sensing," *IEEE Commun. Mag.*, 2021.

[8] A. Guerra, D. Dardari, and P. M. Djurić, "Dynamic radar network of UAVs: A joint navigation and tracking approach," *IEEE Access*, vol. 8, pp. 116454–116469, 2020.

[9] E. Testi, E. Favarelli, and A. Giorgetti, "Reinforcement learning for connected autonomous vehicle localization via UAVs," in *Proc. IEEE Int. Workshop Metrology Agriculture Forestry*, 2020, pp. 13–17.

[10] K. Gu, Y. Wang, and Y. Shen, "Cooperative detection by multi-agent networks in the presence of position uncertainty," *IEEE Trans. Signal Process.*, vol. 68, pp. 5411–5426, 2020.

[11] A. Guerra, D. Dardari, and P. M. Djurić, "Dynamic radar networks of UAVs: A tutorial overview and tracking performance comparison with terrestrial radar networks," *IEEE Veh. Technol. Mag.*, vol. 15, no. 2, pp. 113–120, 2020.

[12] L. Wielandner, E. Leitinger, and K. Witrisal, "Information-criterion-based agent selection for cooperative localization in static networks," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2020, pp. 1–7.

[13] C. Wang et al., "Autonomous navigation of uavs in large-scale complex environments: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2124–2136, 2019.

[14] S. Zhang et al., "Self-aware swarm navigation in autonomous exploration missions," *Proc. IEEE*, vol. 108, no. 7, pp. 1168–1195, 2020.

[15] H. Bayerlein et al., "Multi-UAV path planning for wireless data harvesting with deep reinforcement learning," *arXiv preprint arXiv:2010.12461*, 2020.

[16] M. Theile et al., "UAV path planning using global and local map information with deep reinforcement learning," *arXiv preprint arXiv:2010.06917*, 2020.

[17] O. Esrafilian, R. Gangula, and D. Gesbert, "3D Map-based trajectory design in UAV-aided wireless localization systems," *IEEE Internet of Things J.*, 2020.

[18] H. Bayerlein et al., "UAV path planning for wireless data harvesting: A deep reinforcement learning approach," *arXiv preprint arXiv:2007.00544*, 2020.

[19] M. Lotti et al., "Radio simultaneous localization and mapping in the terahertz band," in *Proc. 25th Int. ITG Workshop on Smart Antennas*, 2021.

[20] S. Ju et al., "Scattering mechanisms and modeling for terahertz wireless communications," in *Proc. IEEE Int. Conf. Commun.*, 2019, pp. 1–7.

[21] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, MIT press, 2018.

[22] R. S. Sutton et al., *Introduction to reinforcement learning*, vol. 135, MIT press Cambridge, 1998.

[23] N. Mafi, F. Abtahi, and I. Fasel, "Information theoretic reward shaping for curiosity driven learning in POMDPs," in *Proc. IEEE Int. Conf. Develop. Learning (ICDL)*, 2011, vol. 2, pp. 1–7.

[24] I. Fasel et al., "Intrinsically motivated information foraging," in *Proc. IEEE 9th Int. Conf. Develop. Learning*, 2010, pp. 101–107.

[25] A. Mariani, A. Giorgetti, and M. Chiani, "Effects of noise power estimation on energy detection for cognitive radio applications," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3410–3420, Dec. 2011.

[26] M. Chiani, "Integral representation and bounds for Marcum Q-function," *Electronics Lett.*, vol. 35, no. 6, pp. 445–446, 1999.

[27] F. Guidi, A. Guerra, and D. Dardari, "Personal mobile radars with millimeter-wave massive arrays for indoor mapping," *IEEE Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1471–1484, 2015.

COLLABORATIVE COMMUNICATIONS BETWEEN A HUMAN AND A RESILIENT SAFETY SUPPORT SYSTEM

Saeideh Samani
NASA Langley Research Center
Hampton, VA, USA
saeideh.samani@nasa.gov

Richard Jessop
Northrop-Grumman
Newport News, VA, USA
rkjessop@gmail.com

Angela Harrivel
NASA Langley Research Center
Hampton, VA, USA
angela.harrivel@nasa.gov

ABSTRACT

Successful introductory UAM integration into the NAS will be contingent on resilient safety systems that support reduced-crew flight operations. In this paper, we present a system that performs three functions: 1) monitors an operator's physiological state; 2) assesses when the operator is experiencing anomalous states; and 3) mitigates risks by a combination of dynamic, context-based unilateral or collaborative dynamic function allocation of operational tasks. The monitoring process receives high data-rate sensor values from eye-tracking and electrocardiogram sensors. The assessment process takes these values and performs a classification that was developed using machine learning algorithms. The mitigation process invokes a collaboration protocol called DFACCto which, based on context, performs vehicle operations that the operator would otherwise routinely execute. This system has been demonstrated in a UAM flight simulator for an operator incapacitation scenario. The methods and initial results as well as relevant UAM and AAM scenarios will be described.

Index Terms— Advanced Air Mobility, Human Monitoring, Communications Between Human and Autonomous Systems, Safe Aviation Transportation

1. INTRODUCTION

The target operational environment for Urban Air Mobility (UAM) envisions a safe, efficient, and economically viable aviation transportation system that will use highly automated aircraft which will operate and transport passengers or cargo at low altitudes within urban and suburban areas. Development of UAM concepts will consider the safety and resilience of the aircraft, the eventual removal of the pilot/operator from the flight deck, the framework for operation, shared access to airspace, infrastructure development, and community engagement. [1]. Advanced Air Mobility (AAM) builds upon the UAM concept by incorporating use cases not specific to operations in urban environments [2]. A significant economic barrier on the introduction of these concepts is the cost of an onboard human vehicle operator. Most current commercial transport flights require two licensed pilots. For UAM vehicle operations to be economically viable, there must be a migration path that simultaneously increases public trust and confidence with UAM operations and reduces the number of operators to one. Eventually, the remaining human operator will be replaced with a fully autonomous flight control system. Once the systems, such as those provided in this paper, are implemented in simulators, research studies can be conducted using them to answer questions regarding the design of increasingly autonomous systems for optimal safety and efficiency [4].

Ensuring safety is an ongoing concern for aviation safety organizations such as the National Aeronautics and Space Administration (NASA) and the Federal Aviation Administration (FAA). The idea of shared mobility services has emerged to not only alleviate demand for parking spaces, but also to reduce the vehicle ownership and environmental impacts [7]. This interest has been reflected in aviation. NASA is working on developing a framework to integrate research called Urban Air Mobility (UAM). This involves an automated electrical Vertical Take-Off and Landing aircraft (e-VTOL) with a capacity of four to six passengers.

Public trust is an important concept and research is needed to understand when people are going to use UAM. Garrow et al. [8] studied this by designing a survey to collect data from 2500 people with high income in different places in the U.S. Another study by Haddad et al. [9] which was based on a survey found some key factors for the public acceptance including trust and safety using in-vehicle cameras and operators, service reliability, and social attitudes.

Additionally, there is a need to study UAM hazard and risk management as well as contingency management; Thipphavong et al. [10] researched some of these factors. In the UAM world, there should be a contingency response manager who is responsible to manage the situation and assist the pilot similar to dispatchers as they support pilots in the current airspace. Furthermore, there is a need to design and develop a framework to manage response to emergencies in UAM operations where the pilot is no longer able to manage the situation with ground-based support.

There are many papers that discuss humans as the main factor of accidents in aviation, and authors express the importance of increasing automation and replacing most of the human tasks with autonomy in the cockpit. On the other hand, the Airline Pilots Association argues and expresses that this is not a valid argument, and well-trained pilots are the core component of aircraft safety [11]. Scerbo et al. described the adaptive automation which requires shared control between the operator and the system [12]. However, there have been discussions about who has the authority over the system and mostly suggest that the operator should have the authority since they are more responsible for the system, and they are able to manage the system in normal or abnormal circumstances. Nevertheless, some researchers believe that the operator might not always perform the best action in different conditions due in part to human performance degradation [13].

Based on multiple studies, fatigue is one of the most important factors of degraded performance as humans interact with autonomy. Therefore, developing technology to increase autonomy and improve the efficiency of human involvement has attracted researchers' interest. Most of the research in human-autonomy teaming is about the need for humans to take over the implementation of a plan and make decisions in critical situations or to return the control initiatives to humans when the automated subtask is complete. Interoperability between human and autonomy called Human-Autonomy Teaming (HAT) [14]. Almost all previous studies agree that humans should be the final authority in HAT systems and always have the ultimate responsibility. However, this could increase human error due to situational awareness issues or ironically a lack of trust in automation [15, 16].

NASA has been researching to find out if a remote pilot can perform similarly to an on-board pilot, and if this remote pilot can assume some of total workload. A report by the U.S. Department of Transportation, FAA, and American Airlines was released to the public in 2004 about in-flight medical incapacitation and impairment of U.S airline pilots during the years 1993–1998 [17]. This study revealed that about one third of surveyed pilots had experienced incapacitation and needed another crewmember to take control of the aircraft. During this 6-year study, about 47 impairment and incapacitation events occurred for pilots with the average age of 43.3 years (range 27–57 years). This report showed that in such cases there definitely is a need for the help of another crew member to take over the pilot's tasks. However, in reduced-crew operations, this back-up pilot would not be available; hence the need for some level of autonomous operations.

Crew State Monitoring System (CSM) software has been designed and developed at the NASA Langley Research Center that uses a broad range of sensors to measure psychophysiological activity of the body in high-fidelity flight simulation studies in real-time. This software supports training methods to reduce accidents and incidents. Attention-related human performance limiting states (AHPLS) is one of the techniques of safety enhancement in the category of "Training for Attention Management." A significant number of aviation accidents involve flight crew distractions due to diverted and channelized attention. For contingency management and detection of psychophysiological states such as channelized attention, diverted attention, and startle/surprise, different physiological sensors have been developed. Multistate classifiers implemented using machine learning and deep learning techniques. Multi-state prediction using this method can identify non-nominal attentional states at rates >80% [2, 18].

This paper describes work to design and develop a platform using a Dynamic Function Allocation Control Collaboration Protocol (DFACCto) which simulates shared or fully autonomous control of the aircraft in case of the pilot's distraction or incapacitation. In this way, automation level could be envisioned to change depending on the current status and workload of the operator. Here, two UAM scenario applications will be briefly described.

The experimental process, user interface, and use cases implemented thus far are discussed in sections 3 and 4.

2. BACKGROUND

Ruff et al. [18] discovered some human factors issues such as operator interaction with automation level and decision process fidelity. They have developed a system called Multi-modal Immersive Intelligent Interface for Remote Operation (MIIIRO) which has the ability for either manual or automatic modes. The design also includes visualization modes to help situational awareness. The operator can update the plan in emergency scenarios when necessary. One such mode is a tactical situation display which is a method of alerting the operator in contingencies. Ruff et al. believe a level of automation is needed to mitigate human factor issues.

Brandt et al. [20] developed a framework for HAT in aviation. They found that HAT displays and automation were preferred for situation awareness, the ability to solve important flight issues, reduced workload, and efficiency. The system called Autonomous Constrained Flight Planner (ACFP) is a recommender system that supports rapid diversion decisions for commercial pilots in non-normal situations [21]. In most of the above human-autonomy teaming studies, intelligence is absent from the automated systems teamed with humans. In our study however, we used machine learning methods, and all of our operations and data analysis is in real-time. We intend to include some level of intelligence or system decision-making using measured operator status.

Psychophysiological sensors have been used in different studies to predict emotional and cognitive states of the body such as workload, attention, and awareness. There are different types of wireless psychophysiological sensors

available for such studies [3–6]. Harrivel et al. studied the prediction of AHPLS by applying psychophysiological sensors and collecting data from a human subject study conducted in a flight simulator. Eye tracking metrics were not included in their classifier inputs due to data drop out caused by gross head motion, although they have mentioned that eye tracking can be used to evaluate crew state awareness. In our study, we have used such methods to predict a crew member's state by applying head-worn eye tracking technology that does not rely on the head being within a pre-determined volume. Communicating information regarding the status of the operator to the autonomous system can help guide the selection of contingency plans for a more graceful transfer of control when the human needs to recover, thus reducing some of the costs associated with HAT.

Levels of automation have increased over time and autonomous systems work with humans as a team to engage in functions such as coordination, task reallocation, and interaction with humans or other systems [22, 23]. However, this concept is not new and has been researched since 1991 [24]. O'Neil et al reviewed multiple papers related to human-autonomy teaming and describes automation in 10 different levels, with level 1 as no computer assistance and all actions and responsibilities are for human, and level 10 wherein the computer decides everything autonomously without any human interaction [25]. They discuss that all the research for human-autonomy teaming is laboratory-based and the studies are based on action or execution simulations with a moderate level of difficulty. Also, they mention that most of the studies address partially—not highly—autonomous systems. In our work however, we begin to design components of a system which is capable of being highly autonomous.

3. DESIGN AND IMPLEMENTATION

We have designed, implemented, demonstrated, and continuously tested a prototype vehicle-based HAT system integrated with a baseline UAM vehicle with custom avionics and control algorithms. This prototype is used to demonstrate the detection of an incapacitated operator which triggers a contingency plan, specifically, a mid-flight redirection to the closest medical facility with a vertipad. The vehicle is a six-passenger, quad-rotor vehicle defined in Silva et al. [26], and the control algorithm is implemented using Python. Important aspects of the methods used include eye-tracking data collection, communication methods, and the Dynamic Function Allocation (DFA) framework.

Recoding eye movements can help us to capture cognitive processes toward accurate state prediction and using machine learning methods, we can analyze eye movement using gaze data. A Tobii Pro glasses 2 eye tracker device (developed by Tobii Technology AB, Danderyd, Sweden) with 50Hz sampling rate and a maximum total system latency of 10 ms is used to record gaze data during our experiments. The eye tracker server was connected to a machine running Tobii Pro Studio. This CSM system is used to capture eye movement data with the Tobii eye tracker in real-time and to record the data for further processing purposes. We receive and analyze eye movement data such as gaze position, gaze direction, and pupil diameter using a Support Vector Machine (SVM) classifier to predict events with our designed event prediction model.

The user interface (UI) is implemented using Qt toolkit and Qt Modeling Language (Qt/QML). We also used voice communication technology since it can be used as a backup communications system based on [27]. In UAM operations, analog voice communications can be used for safety-critical exchanges. In this project, we used User Datagram Protocol (UDP) as provided by a commonly available communications library to transport vehicle data, emergency navigation data, weather data, and other messages.

DFA is a systems engineering/HAT process that seeks to balance the workload for a human operator by distributing operations between the operator and the vehicle's flight operations. DFA operations include those for aviation, navigation, and communication. This framework provides an intuitive command/response interface to vehicle operations that is accessible with the least possible dependency on software engineering expertise. The accessibility includes straightforward development of UIs as well as a well-defined command grammar with an abstracted communications/network protocol. The system also permits transparency of operations which means that it is always explicitly clear who has responsibility for particular operations and provides the positive transfer of control [20]. Because of the nature of human interactions and decision-making processes, the system is asynchronous. Between a command and its response, other responses may be received. Therefore, all messages must be positively identified within a point-to-point conversation.

4. PERFORMANCE EVALUATIONS

4.1. Experimental Setup

To set up the system, we used a machine which had CSM software as well as machine learning methods installed and was connected to an eye tracker. Another machine had DFACCto as well as two software programs that were used to acquire flight plans. These machines plus a UAM simulator machine were on the same network. DFACCto provides integration between the UAM simulator, the CSM software, and autonomous control software. Figure 1 illustrates the system's setup.

We have the participant put on the eye tracker and configure a flight plan. We then initiate the eye-tracking system and then the flight. The UAM simulator has the option of flying in manual or automatic mode depending on the use case scenario we wish to demonstrate. In our

demonstration, we choose the manual mode and have the participant fly the flight plan. If a simulated incapacitation or a distraction is detected, a context-dependent contingency plan is engaged as described below.

Fig. 1. Diagram of the CSM-UAM simulator connections

4.2. Use Cases

4.2.1. Incapacitated Operator

An incapacitated operator is assumed to be a vehicle operator who has lost consciousness or is experiencing microsleep. Such incapacitation may be detected [27] by an eye tracker which can measure various ocular attributes. The prototype system was built such that if the pupil diameter cannot be measured for 5 seconds, DFACCto could execute additional control, an alert, or a diversion. To address an extreme case of incapacitation, the vehicle is diverted from its original flight plan and rerouted to land autonomously at the closest medical facility. To implement system functionality, this scenario was simply demonstrated when closed eyes were detected. Other methods to detect incapacitation should be explored, including possibly pilot stick input behavior [28].

DFACCto extracts navigational information for the closest medical facility from the vehicle and will share that information as well as the vehicle state data with the control software. The control software then computes the navigation and control inputs, e.g., climb, accelerate, turn, and sends those commands to DFACCto. DFACCto can only execute commands that are explicitly provided by the vehicle. These commands are identical to those that the human operator can perform and include the ability to aviate, navigate, and communicate. The details of this protocol implementation are proprietary at this time.

4.2.2. Distracted Operator

A distracted operator is assumed to be a one who appears to be visually distracted from their assigned tasks as detected by an eye tracker. This scenario can occur if the operator is handling an off-nominal vehicle event, inclement weather, or a passenger medical situation. A distracted operator may be detected by using information such as gaze position and gaze direction data recorded from an eye tracker. There were five devices of interest in the simulated vehicle including one primary flight and two secondary system monitors as well as a directional control stick and a novel speed control stick. The model is trained such that if the operator is looking somewhere other than those five devices the event would be predicted as distracted. For this scenario the prototype system was built such that if the gaze data from eye tracker is different from the system's trained data, the SVM would classify that event as distracted. The mitigation control process invokes DFACCto to take autonomous control of the aircraft in case of pilot distraction: the vehicle mode changes to automatic mode and automation executes the existing operation plan.

In this case, the system will perform the mode change command and, if appropriate, hold the command until superseded by the operator. This scenario was demonstrated simply by using an eye tracker, which detected the operator not looking at the devices of interest for at least 5 seconds. Another method has also been presented [29], using heart rate variability, finger plethysmogram amplitude, and perspiration behavior to assess workload. Other methods should be explored, such as multi-modal classifications using galvanic skin response and pre-processed electroencephalography, or measures of autonomous nervous system responses [2–6] to detect an overloaded operator toward the allocation of functions.

5. DISCUSSION AND FUTURE RESEARCH

A real-time system is used with the help of a human-in-the-loop (HITL) air traffic control simulation to explore methods to operate the aircraft safely when the operator is experiencing anomalous states. This system appears to be a promising method which in real-time uses physiological state monitoring to assess when the operator is experiencing anomalous states. For achieving our goal, we combined dynamic, context-based unilateral or collaborative function allocation of operational tasks. This system has been demonstrated end-to-end in a UAM flight simulator for operator incapacitation and distracted operator scenarios. The established foundational methods and successful initial results motivate the development of further use cases relevant to the simulation of UAM and AAM scenarios for HITL research purposes. Finally, the approach should provide a migration path from reduced-crew to fully autonomous flight operations as autonomous capabilities become available and public confidence increases in the use of such technology.

Additional information is needed to operate the aircraft safely and to optimize a positive outcome. Data should be collected to validate detection, execution, and user response times. In the case of incapacitation, this information could include the hospital's location or emergency communication policies, and possibly other data regarding the operator's status using different sensors in addition to the eye tracker.

6. ACKNOWLEDGMENT

This research was funded by the NASA ARMD's System Wide Safety Program and performed by NASA Langley Research Center's Crew Systems and Aviation Operations Branch. One or more aspects of this disclosure are patent pending and covered in at least: U.S. Pat. App. No. 15/908,026.

7. REFERENCES

[1] https://www.faa.gov/uas/advanced_operations/urban_air_mobility

[2] Harrivel, A.R., Liles, C., Stephens, C.L., Ellis, K.K., Prinzel, L.J. and Pope, A.T., 2016. Psychophysiological sensing and state classification for attention management in commercial aviation. In AIAA Infotech@ Aerospace (p. 1490).

[3] Novak, D., Mihelj, M., Munih, M., "A survey of methods for data fusion and system adaption using autonomic nervous system responses in physiological computing," Interacting with Computers, Vol. 24, 2012, pp. 154-172.

[4] Pope, A. T., Bogart, E. H., Bartolome, E. S., "Biocybernetic system evaluates indices of operator engagement in automated task," Biological Psychology, Vol. 40, 1995, pp. 187-195.

[5] Wilson G. F. and Russell, C. A., "Real-time assessment of mental workload using psychophysiological measures and artificial neural networks," Human Factors, Vol. 45, No. 4, 2003, pp. 635-643.

[6] Fairclough, S., and Gilleade, K., "Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation," International Journal of Autonomous and Adaptive Communications Systems, Vol. 6, No. 1, 2013, pp. 63–79.

[7] Baptista, P., Melo, S., & Rolim, C. (2014). Energy, environmental and mobility impacts of car-sharing systems. Empirical results from Lisbon, Portugal. Procedia-Social and Behavioral Sciences, 111, 28-37.

[8] Garrow, L. A., German, B., Mokhtarian, P., Daskilewicz, M., Douthat, T. H., & Binder, R. (2018). If you fly it, will commuters come? A survey to model demand for e-VTOL urban air trips. In 2018 Aviation Technology, Integration, and Operations Conference (p. 2882).

[9] Michelmann, J., Straubinger, A., Becker, A., Al Haddad, C., Plötner, K.O. and Hornung, M., 2020. Urban Air Mobility 2030+: Pathways for UAM-A Scenario-Based Analysis. In Deutscher Luft-und aumfahrtkongress 2020.

[10] Thipphavong, D. P., Apaza, R., Barmore, B., Battiste, V., Burian, B., Dao, Q., Feary, M., Go, S., Goodrich, K. H., Homola, J. et al. (2018). Urban air mobility airspace integration concepts and considerations. In 2018 Aviation Technology, Integration, and Operations Conference (p. 3676).

[11] Air Line Pilots Association, "Airline pilots Association white paper on Unmanned Aircraft Systems," April 2011

[12] Scerbo, M.W., 2006. Adaptive Automation.

[13] Wiener, E.L., 1989. Human factors of advanced technology (glass cockpit) transport aircraft.

[14] McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. Human Factors: The Journal of the Human Factors and Ergonomics Society, 60, 262–273. https:// doi. org/ 10. 1177/ 0018720817743223

[15] Goodrich, M. A. 2013. "Multitasking and Multi-Robot Management." In Oxford Handbook of Cognitive Engineering, edited by J. D. Lee and A. Kirklik, 379–394. New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780199757183.013.0025.

[16] Barnes, M. J., J. Y. C. Chen, and F. Jentsch. 2015. "Designing for Mixed-Initiative Interactions between Human and Autonomous Systems in Complex Environments." 2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Hong Kong: IEEE. 1386–1390. doi: 10.1109/SMC.2015.246.

[17] DeJohn, C.A., Wolbrink, A.M. and Larcher, J.G., 2004. In-flight medical incapacitation and impairment of US airline pilots: 1993 to 1998 (No. DOT/FAA/AM-04/16). FEDERAL AVIATION ADMINISTRATION OKLAHOMA CITY OK CIVIL AEROMEDICAL INST.

[18] Terwilliger, P., Sarle, J., Walker, S., Harrivel, A. A ResNet Autoencoder Approach for Time Series Classification of Cognitive State, MODSIM World 2020, Paper No. 0053, Norfolk VA, May 5-7, 2020

[19] H. A. Ruff, S. Narayanan, and M. H. Draper, "Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles," Presence: Teleoperators & Virtual Environments, Special Issue on Virtual Environments and Mobile Robots: Control, Simulation, and Robot Pilot Training, vol. 11, pp. 335-351, Aug. 2002.

[20] Brandt, S.L., Lachter, J., Russell, R. and Shively, R.J., 2017, July. A human-autonomy teaming approach for a flight-following task. In International Conference on Applied Human Factors and Ergonomics (pp. 12-22). Springer, Cham.

[21] Dao, A-Q., Koltai, K., Cals, S.D., Brandt, S.L., Lachter, J., Matessa, M., Smith, D., Battiste, V., Johnson, W.W.: Evaluation of a Recommender System for Single-Pilot Operations. Procedia Manufacturing. 3, 3070--3077 (2015)

[22] Shannon, C. J., Horney, D. C., Jackson, K. F., & How, J. P. (2017). Human-autonomy teaming using flexible human performance models: An initial pilot study. In Advances in human factors in robots and unmanned systems (pp. 211–224). Springer. https:// doi. org/ 10. 1007/ 978- 3- 319-41959- 6_ 18.

[23] Cooke, N., Demir, M., & McNeese, N. (2016). Synthetic teammates as team players: Coordination of human and synthetic teammates (Report No. RE2016844 01). Cognitive Engineering Research Institute Mesa United States.

[24] Malin, J. T., Schreckenghost, D. L., Woods, D. D., Potter, S. S., Johannesen, L., Holloway, M., & Forbus, K. D. (1991). Making intelligent systems team players: Case studies and design issues. Volume 1: Human-computer interaction design. NASA Johnson Space Center.

[25] O'Neill, T., McNeese, N., Barron, A. and Schelble, B., 2020. Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. Human Factors, p.0018720820960865.

[26] Silva, C., Johnson, W.R., Solis, E., Patterson, M.D. and Antcliff, K.R., 2018. VTOL urban air mobility concept vehicles for technology development. In 2018 Aviation Technology, Integration, and Operations Conference (p. 3847).

[27] Golz, M., Sommer, D., Chen, M. et al. Feature Fusion for the Detection of Microsleep Events. J VLSI Sign Process Syst Sign Im 49, 329–342 (2007). https://doi.org/10.1007/s11265-007-0083-4

[28] Trujillo, A.C., Gregory, I.M., 2016. Wetware, Hardware, or Software Incapacitation: Observational Methods to Determine When Autonomy Should Assume Control. In AIAA AVIATION 2014, Atlanta, GA

[29] Miyake S. Multivariate workload evaluation combining physiological and subjective measures. Int J Psychophysiol. 2001 Apr;40(3):233-8. doi: 10.1016/s0167-8760(00)00191-4. PMID: 11228350."

LANE CHANGING USING MULTI-AGENT DQN

Karthikeyan Nagarajan, Zhong Yi

Moovita Pte Ltd
Singapore

ABSTRACT

This study explores the feasibility of autonomous lane changing using a novel approach of multi-agent Deep Q-Network. Most existing algorithms that use Deep Reinforcement Learning adopt a single-agent approach, with the assumption of only ego-agent changing lanes. We argue that such an approach is merely a simplification of the real-world without considering multi-agent negotiations. In this work, we model the lane changing problem as a multi-agent system and develop a decision-making policy using Deep Q-Network. We address the non-stationarity problem caused by our multi-agent setup which includes an Experience Replay. While prior research recommends avoiding the Experience Replay under such conditions, we report for the first time that an Experience Replay can help yield a robust negotiation policy in our lane changing experiment, without impairing the training of the Deep Q-Network. We show that our approach enables the model to learn negotiating-behaviors like overtaking, yielding, lane interchanging, and lane merging.

Index Terms— Autonomous Driving, Deep Reinforcement Learning, Multi-agent Reinforcement Learning

1. INTRODUCTION

In recent years, advancements in machine learning have contributed majorly to the progress of autonomous driving. Though there is a great progress in the sensor system, the decision-making part of the autonomous driving is highly challenging and is still under active research. One such challenge is to develop an intelligent decision-making system which aids in changing lanes based on traffic conditions. With the recent success of the single-agent Deep Reinforcement Learning (DRL), many researchers have proposed methods [1, 2, 3, 4] to develop a policy for lane changing behavior using DRL. DRL-based methods have proven to be more effective compared to the traditional approaches, due to their ability to adapt the rules automatically based on varying traffic situations. However, most of the DRL-based works [1, 3, 5] consider the problem only from a limited perspective, *i.e.*, looking for gaps and changing lanes safely while driving straight. There is no notion of a goal like reaching a target lane before an intersection, a high-way exit, and so on. In reality, the driver strategically uses the lane changing action to accomplish his/her long term mission of navigating his/her vehicle to the final destination. For such a lane changing action, the short term goal of the driver is to reach the target lane within a certain distance as quickly as possible by negotiating through surrounding traffic. The maneuver should be optimal without unnecessary lane changes and should be safe without any collisions. One example of such a goal-based approach is described in [4], in which the goal of the ego-agent is to perform a specific highway exit. Despite showing successful results, most DRL-based works [1, 3, 4] model the problem as a single-agent system and assume that all the agents except ego-agent cannot change lanes. As a result, the developed policy takes unsafe actions assuming that the surrounding agents will drive only straight, which is a simplification of the real-world. The single-agent approach also limits the resource efficiency because only the ego-agent is used to explore, collect experiences and learn the policy. In view of the above, a better formulation would allow all agents to learn to negotiate and cooperate simultaneously, and to consider each other's actions when planning their strategy for their individual goals.

In this paper, we model the lane changing problem as a multi-agent system and develop a decision-making policy using Deep Q-Network (DQN) [6]. The goal of each agent is set to reach its target lane within a fixed distance as quickly as possible while in a safe and efficient manner. The experiences from all the agents are collected simultaneously to learn one common policy. This approach falls under the category called decentralized learning using parameter sharing [7]. However, our setup introduces the problem of non-stationarity [8, 9], which restricts the direct use of Experience Replay [6, 10] along with DQN. Contrary to the common belief that Experience Replay should not be used together with DQN, we demonstrate that it can improve the efficacy and help develop a negotiation policy in our non-stationary environment. We explain and discuss the underlying rationale in Section 2.

2. RESEARCH BACKGROUND

Recently, the use of deep neural networks has dramatically improved the scalability of single-agent Reinforcement Learning (RL). One element key to the success of such ap-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

proaches is the reliance on Experience Replay. Experience replay not only helps to stabilize the training of single-agent DQN, but it also improves sample efficiency by repeatedly reusing experience tuples. However, according to the research [5, 7, 11, 12, 13], it might not provide similar benefits in our multi-agent DQN setup due to the non-stationarity of the environment. Under such conditions, the replay buffer stores the outdated experiences which no longer reflect the current dynamics in which it is learning. Since the replay buffer constantly confuses the agent with obsolete experiences, Experience Replay impairs the training of DQN. In the view of the above, the prior research states that Experience Replay is incompatible with DQN in a non-stationary environment.

However, there are a few research works [5, 7, 11, 12, 13] which propose solutions to mitigate the problem of incompatibility of Experience Replay with DQN in a non-stationary environment. For example, [12] studied the effect of replay buffer size on the trained agent's performance and suggested to limit experience replay buffer to a short, recent buffer. In [13], the replay buffer is completely disabled while learning the policy using DQN. However, the above workarounds limit the sample efficiency because we are limiting the use of experience replay buffer. Other studies [7, 11] propose alternative methods to use Experience Replay without affecting the sample efficiency. But, these methods increase the computation overhead and the complexity of the algorithm.

On the other hand, policy-based RL methods directly map the states to the actions which maximize the expected cumulative reward. Unlike value-based RL methods, e.g. DQN, policy-based methods do not require Experience Replay as they always depend on the experiences generated using the current policy. One such policy-based RL work [14] introduces a technique called *opponent sampling* to improve the performance of PPO in a multi-agent self-play. Instead of using only the current policy version, the technique also samples the older versions of the policy for the opponent during the self-play training. The authors in [14] show that the technique helps to develop a robust policy by avoiding over-fitting to the most recent opponent's policy. Inspired by such a technique, we are interested in checking if training a multi-agent DQN with a large experience replay buffer can prevent the over-fitting of Q-network to the most recent dynamics of the environment, *i.e.*, the most recently evolved policy. Thus, we intuitively explore the effectiveness of Experience Replay to yield a robust policy in our non-stationary environment.

3. METHODOLOGY

3.1. Environment

Carla simulator is chosen as a simulation platform to conduct our experiments [15]. Considering the real-time verification of the algorithm in our restricted environment in the future, a straight one-way road with three lanes is considered in our work. In our experiments, we vary the number of agents n across the episodes to ensure the scalability of the algorithm. Though our setup is an open system [16], the maximum number of agents that are spawned in an episode is limited to 4 due to the computational speed limitations. Based on our test vehicle's capability, the maximum speed v_{max} is randomly selected up to 5 ms^{-1}. The Intelligent Driver Model (IDM) is used as the Adaptive Cruise Control module which defines the target speed and avoids forward-collision [17]. An episode is terminated when a collision happens or when all the agents cross the finish line or timeout. The finish line refers to the horizontal line up of the target flags as depicted in the snapshots in Section 4. The maximum distance to the finish line is $115m$. The time horizon T is equal to the entire duration of the episode and the maximum duration of an episode is $75s$. The agents are spawned at random locations before the finish line across the three lanes such that the longitudinal distance between any two agents is atleast 10 m.

Table 1. Local Observation Space

Input	*Description*
ob_1	Presence of a left lane to change lanes
ob_2	Presence of a right lane to change lanes
ob_3	Ego-agent's speed
ob_4	Ego-agent's longitudinal distance to target
ob_5	Relative target lane id
ob_{3i+1}	Surrounding agent's relative longitudinal distance
ob_{3i+2}	Surrounding agent's relative longitudinal velocity
ob_{3i+3}	Surrounding agent's relative lane id

3.2. Q-Network

Table 1 represents the details of the local observation space from the perspective of each agent. For each agent, the first five elements in its local observation space correspond to the information about itself and the remaining elements ob_{3i+1}, ob_{3i+2} and ob_{3i+3} corresponds to the relative information of the surrounding agents, where $i = 1, \ldots, 10$. To demonstrate scalability later in Section 4.3, the observation space is designed to house up to 10 (arbitrarily chosen limit) surrounding agents'information. Since we spawn only up to 4 agents in an episode during training, the observation space is partially filled, *i.e.*, $1 \leq i \leq 3$. In a real case scenario, the presence of valid lanes on the either side of the ego-vehicle ob_1 and ob_2 are provided by a high-definition map. As shown in Table 2, the action space is discrete and is the same for all agents. a_2 ensures the agent stays in the current lane and the set speed of its low-level speed controller is dictated by its IDM. When the agent selects a_4, the set speed is set to half of the current speed of the agent overriding the value set by the IDM.

The complete model architecture is presented in Figure 1. Similar to the network architecture of [13], the first convolutional layer has 64 filters, each of size 3 and stride 3 so

Table 2. Action Space

Action	*Description*
a_1	Change lane to left
a_2	Stay in the current lane
a_3	Change lane to right
a_4	Stay in the current lane and slow down

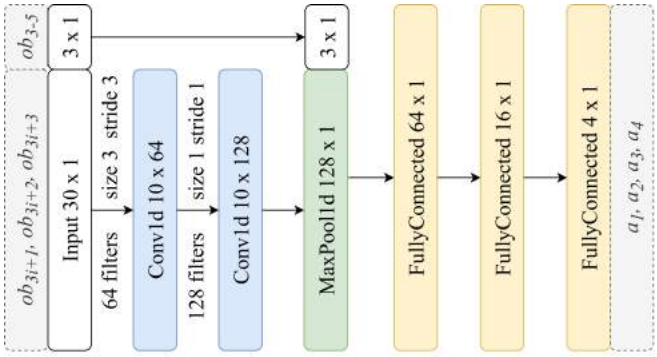


Fig. 1. Network architecture of our Deep Q-Network.

that each filter processes only one surrounding agents information, i.e., ob_{3i+1} , ob_{3i+2} and ob_{3i+3}, at a time and gives one output value for each surrounding agent. Similarly, the second convolutional layer has 128 filters, each of size 1 and stride 1. This aggregates the knowledge about each surrounding agent in every row of 10x128 output. The max-pooling layer consolidates the 10x128 aggregated output into 1x128 values. So, irrespective of the number of surrounding agents, the max-pooling layer consolidates the processed information of all the surrounding agents into 1x128 values. Such an arrangement of convolutional layers and the max-pooling layer not only makes the model translational invariant to the information within the observation space but also facilitates scalability. To enhance exploration efficiency, we have also adopted the technique called Q-masking [4] to mask a_1 and a_3 actions based on the availability of lanes ob_1 and ob_2.

Reward shaping which includes credit assignment is a very important aspect as it enables the agents to learn a cooperative behavior. Our reward shaping is as follows: The agent which is responsible for a collision is punished with a negative reward of -1. To avoid unnecessary lane changes and slow down, any lane change action or *slow down* action is assigned with an immediate reward of -0.05 and -0.03 respectively. If an agent reaches the target lane, it is rewarded with a positive value inversely proportional to the time taken to reach the target lane. This encourages the agent to reach the target lane as fast as possible. If an agent reaches the wrong lane, it is given a negative reward which is proportional to the relative lateral distance from the target lane.

3.3. Training Details

The DQN training is carried out for 15k episodes with only one agent performing ϵ-greedy exploration in an episode. The agents take action at an interval of $1s$ and they do not take decisions during a lane change or *slow down* action. The experiments (*ER102*, *ER104*, *ER106* and *ER108*) are carried out with the same set of hyperparameters except varying replay buffer size M (10^2, 10^4, 10^6 and 10^8) to study the effect of replay buffer size on the policy. For the sake of benchmarking against the single-agent lane changing approaches, *ER106* is repeated with an assumption that the only one random agent in an episode can change lanes and this experiment is named as *S-ER106*.

4. RESULTS

It is observed that there is no standard method of evaluating lane changing behavior since each prior work uses a different set of metrics. With the intention to standardize the evaluation metrics, we use four informative metrics namely, *success%*, *failure%*, *timeout%* and *accident%* apart from one consolidated *score*. The *score* is defined as the mean of the returns of all the agents in an episode. All the agents that have not reached the finish line are termed *active* in an episode. If an active agent reaches the target lane at the finish line, its mission is *success* otherwise, *failure*. If an active agent collides with another active agent, it is considered as an *accident* for the responsible agent. The evaluation of the behavior of the remaining active agents are ignored and the episode is terminated. When the episode times out, it is a *timeout* for all the active agents.

4.1. Quantitative Results

The *ER102*, *ER104*, *ER106*, *ER108* and *S-ER106* are tested with the same set of 2500 randomly generated scenarios. For each test run, the average values of the scores and other metrics are calculated across all the scenarios. From Table 3, the performance degrades as the replay buffer size is decreased from 10^8 to 10^2. There is only a slight difference in performance between *ER106* and *ER108*. The model learns well when the replay buffer size is large. Increasing the buffer size not only decreases the *accident%* but also decreases the *timeout%*. This indicates that the large buffer size does not make the policy very conservative to avoid the accidents. Instead, it results in robust and decently conservative policy. In contrast to the claims of [12] and [13], limiting the replay buffer results in sample inefficiency and thus leading to over-fitting of the Q-network in our case. Moreover, there is no need for any additional methods as specified in [5, 7, 11] to handle the non-stationarity in our environment. Our performance results cannot be benchmarked directly against the lane changing prior works because the evaluation method varies in each case. However, the higher value of *average*

Table 3. Performance Analysis

Name	Success %	Failure %	Accident %	Timeout %	Avg score
S-ER106	94.76	**0.43**	2.83	1.99	0.6641
ER102	89.45	1.60	1.97	6.98	0.6017
ER104	91.74	2.36	0.90	5.01	0.6706
ER106	**97.32**	1.20	**0.50**	0.98	**0.7299**
ER108	96.95	1.85	0.54	**0.66**	0.7238

Table 4. Scalability Analysis

n	Success %	Failure %	Accident %	Timeout %	Avg score
4	97.32	1.20	0.50	0.98	0.7299
6	93.13	3.24	1.99	1.64	0.6300
8	87.97	6.19	2.99	2.85	0.5150
10	82.58	8.06	4.15	5.21	0.4041

score of *ER106*, when compared with that of *S-ER106*, indicates that our multi-agent approach performs better than the single-agent approaches used in the prior work [1] [3] [4].

4.2. Qualitative Results

Apart from the random scenarios testing, the best performer *ER106* is also tested with 36 basic handcrafted scenarios. These scenarios are created to analyze how two agents negotiate during yielding, overtaking, lane interchanging and merging situations. Some negotiating-behaviors exhibited in the handcrafted scenarios are illustrated with the snapshots in Figure 2(a) and Figure 2(b). This proves that the decentralized learning by parameter sharing with proper credit assignment can result in a cooperative behavior without explicit communication between agents [7].

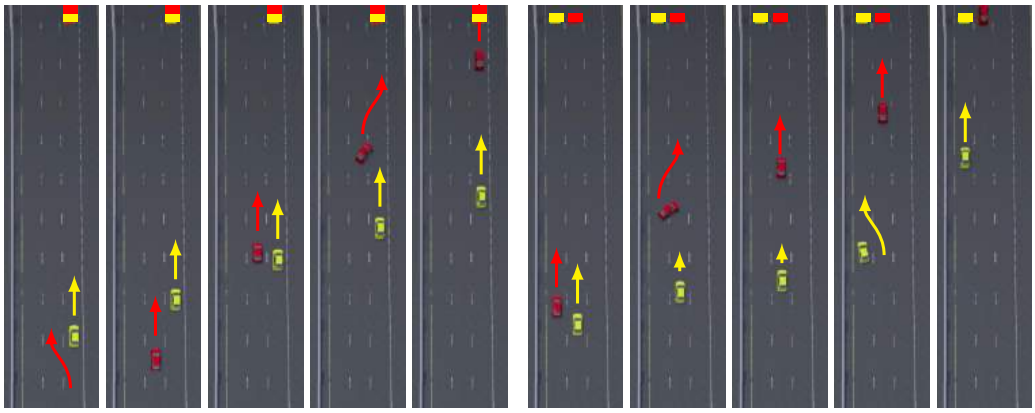
4.3. Scalability

Though [3] utilizes the convolution to achieve translational invariance of the observation space, there is no demonstration of the scalability of their algorithm. In our work, we conduct experiments to analyze the scalability of our approach. Though *ER106* model is trained only with a maximum of 4 agents, we decide to test its endurance with an increased number of agents (6,8, and 10) without any additional training. The maximum duration of the episode is increased to $120s$ to provide sufficient time for all the agents to reach their respective target lanes. According to Table 4, though there is a decline in the performance with an increase in the number of agents, the model performs reasonably well without even being trained in such a scaled-up environment. So, the model learns itself to determine which all agents information are critical among all the information present in the observation space. Our approach not only demonstrates translational invariance of the observation space but also can be easily scaled up with further training.

5. DISCUSSION

Why does the assumption of 'only ego-agent changing lanes' in a single-agent approach reduce the performance? According to Table 3, the high value of *accident%* of *S-ER106* indicates that the single-agent approach used along with the assumption leads to less safe behavior. To match the real-world settings as much as possible, we decided to test the *S-ER106* model in the multi-agent environment. Under such conditions, all the agents can change lanes using the *S-ER106* policy in contrast to the training environment. It is observed that an agent does not expect the other agents to *slow down* for yielding or merging. This is because all agents except ego-agent choose only the action a_2 during *S-ER106* training. So, during testing, each agent changes lane aggressively without expecting the surrounding agents to *slow down* and change lanes. This leads to more collisions and accounts for the high value of *accident%* of *S-ER106*. The above result proves that the assumption used in the prior works [1, 3, 4] reduces the performance of the lane changing policy.

Why does large replay buffer improve the performance instead of impairing it? Though all the agents learn one common policy in our environment, the policy evolves during the training. Under such conditions, a large replay buffer stores experiences that represent the varying dynamics of the environment and thus, it is expected to impair the DQN training. By contrast, our experimental results not only demonstrate that the DQN learning converges with a large replay buffer but also show that the replay buffer helps to develop a better policy in a non-stationary environment. Hence, we empirically deduce that, though the state transition function and the reward function change, the large replay buffer averages out the perceived functions towards the end of the training. For example, at the early stage of training, the lane changing policy is aggressive and causes unnecessary lane changes. After some episodes, the model becomes very conservative and chooses more of *slow down* action. Though different behaviors are exhibited by the surrounding agents during the evolution of the policy, the algorithm explores to find an optimal strategy under all such circumstances and stores the related experiences in one large replay buffer. Hence, after training the Q-network with all such experiences, the resultant q-values corresponds to the averaged out versions of the state transition function and the reward function, and thus resulting in a more robust and generalized policy. From another perspective, since large replay buffer helps to retain experiences from a wide range of training scenarios, increasing the buffer size avoids over-fitting of the Q-Network. A more detailed investigation is left for future work.



(a) Overtaking Scenario (b) Interchanging Scenario

Fig. 2. (a) Snapshots show how the red agent ($v_{max} = 5ms^{-1}$) overtakes the slow-moving yellow agent ($v_{max} = 1.5ms^{-1}$) in the front. (b) Snapshots show how the red agent ($v_{max} = 5ms^{-1}$) and the yellow agent ($v_{max} = 5ms^{-1}$) negotiate without explicit communication and interchange their lanes successfully. The yellow agent decides to slow down while allowing the red agent to move forward and change lane. Best viewed in color prints.

6. CONCLUSION

We have proposed a pragmatic multi-agent DQN approach to develop a decision-making policy for autonomous lane changing. Our results show that the trained model performs well with the success rate above 97% and exhibits some negotiating-behaviors without explicit communication. In contrast to the common belief that a larger replay buffer impairs DQN learning, we have found that it yields a robust policy and improves the performance by avoiding over-fitting of the Q-Network. Hence, we deduce that the incompatibility of Experience Replay with DQN cannot be generalized for all non-stationary environments. Our research work opens up the opportunities for researchers in the field of Deep Reinforcement Learning to explore the influence of Experience Replay on the DQN-based policy in a non-stationary environment.

References

[1] Chen Chen, Jun Qian, Hengshuai Yao, Jun Luo, Hongbo Zhang, and Wulong Liu, "Towards comprehensive maneuver decisions for lane change using reinforcement learning," in *32st Conference on Neural Information Processing Systems (NIPS 2018)*, 2018.

[2] Junjie Wang, Qichao Zhang, Dongbin Zhao, and Yaran Chen, "Lane change decision-making through deep reinforcement learning with rule-based constraints," *arXiv preprint arXiv:1904.00231*, 2019.

[3] Carl-Johan Hoel, Krister Wolff, and Leo Laine, "Automated speed and lane change decision making using deep reinforcement learning," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 2148–2155.

[4] Mustafa Mukadam, Akansel Cosgun, Alireza Nakhaei, and Kikuo Fujimura, "Tactical decision making for lane changing with deep reinforcement learning," in *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 2017.

[5] Safa Cicek, Alireza Nakhaei, Stefano Soatto, and Kikuo Fujimura, "Marl-pps: Multi-agent reinforcement learning with periodic parameter sharing," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, Richland, SC, 2019, AAMAS 19, p. 18831885, International Foundation for Autonomous Agents and Multiagent Systems.

[6] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529, 2015.

[7] Jayesh K. Gupta, Maxim Egorov, and Mykel Kochenderfer, "Cooperative multi-agent control using deep reinforcement learning," in *Autonomous Agents and Multiagent Systems*. 2017, pp. 66–83, Springer International Publishing.

[8] Pablo Hernandez-Leal, Bilal Kartal, and Matthew E. Taylor, "A survey and critique of multiagent deep reinforcement learning," *Autonomous Agents and Multi-Agent Systems*, vol. 33, no. 6, pp. 750797, Oct 2019.

[9] Georgios Papoudakis, Filippos Christianos, Arrasy Rahman, and Stefano V. Albrecht, "Dealing with non-stationarity in multi-agent deep reinforcement learning," 2019.

[10] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver, "Prioritized experience replay," in *International Conference on Learning Representations*, Puerto Rico, 2016.

[11] Jakob Foerster, Nantas Nardelli, Gregory Farquhar, Triantafyllos Afouras, Philip H. S. Torr, Pushmeet Kohli, and Shimon Whiteson, "Stabilising experience replay for deep multi-agent reinforcement learning," in *Proceedings of the 34th International Conference on Machine Learning*, Doina Precup and Yee Whye Teh, Eds., International Convention Centre, Sydney, Australia, 06–11 Aug 2017, vol. 70 of *Proceedings of Machine Learning Research*, pp. 1146–1155, PMLR.

[12] Joel Z. Leibo, Vinicius Zambaldi, Marc Lanctot, Janusz Marecki, and Thore Graepel, "Multi-agent reinforcement learning in sequential social dilemmas," in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, Richland, SC, 2017, AAMAS 17, p. 464473, International Foundation for Autonomous Agents and Multiagent Systems.

[13] Jakob N. Foerster, Yannis M. Assael, Nando de Freitas, and Shimon Whiteson, "Learning to communicate with deep multi-agent reinforcement learning," in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, 2016, NIPS16, p. 21452153, Curran Associates Inc.

[14] Trapit Bansal, Jakub Pachocki, Szymon Sidor, Ilya Sutskever, and Igor Mordatch, "Emergent complexity via multi-agent competition," in *International Conference on Learning Representations*, 2018.

[15] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun, "Carla: An open urban driving simulator," 2017.

[16] Muthukumaran Chandrasekaran, Adam Eck, Prashant Doshi, and Leenkiat Soh, "Individual planning in open and typed agent systems," in *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*, Arlington, Virginia, USA, 2016, UAI16, p. 8291, AUAI Press.

[17] Martin Treiber, Ansgar Hennecke, and Dirk Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical Review E*, vol. 62, pp. 1805–1824, 02 2000.

DATA-DRIVEN PUMP SCHEDULING FOR COST MINIMIZATION IN WATER NETWORKS

Jyotirmoy Bhardwaj, Joshin Krishnan, Baltasar Beferull-Lozano

WISENET Center, Department of Information & Communication Technology,
University of Agder, Grimstad, Norway
{jyotirmoy.bhardwaj, joshin.krishnan, baltasar.beferull}@uia.no

ABSTRACT

Pumps consume a significant amount of energy in a water distribution network (WDN). With the emergence of dynamic energy cost, the pump scheduling as per user demand is a computationally challenging task. Computing the decision variables of pump scheduling relies over mixed integer optimization (MIO) formulations. However, MIO formulations are NP-hard in general and solving such problems is inefficient in terms of computation time and memory. Moreover, the computational complexity of solving such MIO formulations increases exponentially with the size of the WDN. As an alternative, we propose a data-driven approach to estimate the decision variables of pump scheduling using deep neural networks (DNN). We evaluate the performance of our trained DNN relative to a state-of-the-art MIO solver, and conclude that our DNN based approach can be used to minimize the pump switching and cost incurred due to dynamic energy in a given WDN with much lower complexity.

***Index Terms*—Pump scheduling, mixed-integer formulation, deep neural networks, dynamic energy cost, water-energy nexus.**

I. INTRODUCTION

Pump scheduling is an integral part of water distribution network (WDN) management. As per energy statistics of US, WDNs and treatment plants consume approximately 4% of total produced energy [1]. Pumps consume a significant amount of energy, and optimal pump scheduling can save the energy consumption by 10%-20% [2]. The empirical flow and energy constraints imposed by WDN are non-convex, and the decision variables (pumps and valve control) are binary [3]. Hence, most approaches construct the pump scheduling problem as a *mixed-integer optimization* (MIO) problem with an objective to minimize the energy cost of water dispatch to consumers [4‡ [5]. In general, MIO is an $\mathcal{NP}$-hard non-convex problem, and computing the decision variables of such problems takes substantial memory and computation time with the current approaches [6]. Moreover, the computational complexity further increases with the ever growing expansion of WDNs.

In contrast, we propose a data-driven approach, in which the decision variables of MIO formulations are learned from the data set of a WDN, and bypass the need of any MIO solver. This approach is motivated by the observation that a feed-forward deep neural network (DNN) can estimate the decision variables of MIO problems with high accuracy. Such data-driven approaches exploit the repetitive patterns of problem instances, and reduce the MIO formulations to a neural network prediction, and once the DNN model is learned, it can speedup the computation time to solve MIO problem for new problem instances [7].

This paper has two major contributions. First, we propose a MIO framework for pump scheduling in a WDN aiming at minimizing the energy cost, given the time-ahead dynamic electric energy prize. The proposed optimization framework also ensures that the WDN parameters are within the admissible operating range, which is a difficult task to achieve by a human operator especially for large-scaled WDNs.The second contribution focuses on a data-driven pump scheduling strategy based on training a feed forward DNN. We solve an offline MIO problem for given network constraints using a MIO solver, obtain the values of decision variables, and train a feed-forward DNN using those values. We benchmark the performance of DNN against the state-of-the-art MIO solver Gurobi [8] using experiments conducted over synthetic data sets, which shows that our approach bypass the need of a solver.

The rest of the paper is structured as follows. Section II formulates an MIO problem in WDN with an objective to minimize the pump switching and cost incurred due to dynamic energy. Section III proposes a model-free data-driven approach for estimation of decision variables in the WDN. Section IV and Section V respectively, present the experimental results and conclusion.

II. MIO FOR PUMP SCHEDULING

Model: Consider a WDN modeled as a directed graph $\mathcal{G}$:=$(\mathcal{N}, \mathcal{P})$, where $\mathcal{N}$ and $\mathcal{P}$ denote the sets of nodes and edges (pipes) of the WDN, respectively. A typical WDN

This work was supported in part by IKTPLUSS funded Project "Data-driven cyber-physical networked systems for autonomous cognitive control and adaptive learning in industrial urban water environments (INDURB)", led by WISENET Center, University of Agder, Norway.

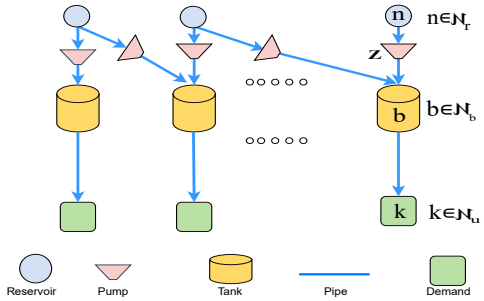


Fig. 1: Schematic of a WDN.

mainly includes the following types of nodes: **a)** reservoirs, the primary source of water; **b)** tanks, the nodes that store water from the reservoirs; and **c)** users, the points of water demands. Let $\mathcal{N}_r \subset \mathcal{N}$ and $\mathcal{N}_b \subset \mathcal{N}$ be respectively the subsets of nodes corresponding to the reservoirs and the tanks, and $\mathcal{N}_u := \mathcal{N} \setminus \mathcal{N}_r \cup \mathcal{N}_b$ be the subsets of nodes denoting the users. An edge between node n and node k of the graph is denoted by the directed pair $(n,k) \in \mathcal{P}$; and let $\mathcal{P}_a \subset \mathcal{P}$ be the subset of edges that host the pumps. In our model, we assume that the pumps are deployed only between the reservoirs and the tanks, and the water flow from tanks to users is generated by means of gravity; hence, $(n,b) \in \mathcal{P}_a \implies n \in \mathcal{N}_r, b \in \mathcal{N}_b$ as shown in Fig. 1.

Flow Conservation: Let q_t^k be the net water outflow rate from the node $k \in \mathcal{N}$ at the time index t, where $t = 1, 2, \ldots, T$. The water flow from node n to node k at time t is denoted as $q_t^{nk} \geq 0$, with $q_t^{nk} = 0$ when there is no flow. Further, we assume a directed graph, meaning that $q_t^{kn} = 0$ when $q_t^{nk} \geq 0$. Considering the flow conservation, the net water outflow at node k, can be written as

$$q_t^k = \sum_{(k,n)\in\mathcal{P}} q_t^{kn} - \sum_{(n,k)\in\mathcal{P}} q_t^{nk}, \quad \forall t \in \mathcal{T}, \tag{1}$$

where $\mathcal{T} = \{0, 1, \ldots, T-1\}$ is the set of time indices. Further, the water flow is constrained as

$$0 \leq q_t^{nk} \leq q^{\max}, \quad \forall\ (n,k) \in \mathcal{P}, \quad t \in \mathcal{T}, \tag{2}$$

where $q^{\max}$ is the maximum possible flow on any edge, as the flow is constrained by physical properties of the pipe such as length, diameter and friction coefficient. This capacity bound can be obtained from empirical head-loss equations such as Hazen-Williams or Darcy-Weisbach equations [9].

Tank formulation: We assume that each tank $b \in \mathcal{N}_b$ receives water from the reservoirs $n \in \mathcal{N}_r$ through the pump-hosting edges $(n,b) \in \mathcal{P}_a$ and they supply water to the rest of the network. The dynamics of the water storage tank can be modelled as

$$v_{t+1}^b = v_t^b + \tau \left(\sum_{(n,b)\in\mathcal{P}_a} q_t^{nb} - \sum_{(b,n)\in\mathcal{P}\setminus\mathcal{P}_a} q_t^{bn} \right), \tag{3}$$

where v_t^b is the amount of water stored in tank b at time index t and $\tau > 0$ is the sampling time interval. We also bound v_t^b to avoid the overflow of the tank and to meet any unexpected user demand:

$$v^{\min} \leq v_t^b \leq v^{\max}, \ b \in \mathcal{N}_b, \ \forall t \in \mathcal{T}. \tag{4}$$

Where, $v^{\min}$ and $v^{\max}$ is the minimum and maximum water volume in a tank respectively.

Pump Formulation: We introduce a binary variable $z_t^{nb} \in \{0,1\}$ to denote the pump switching in the edge $(n,b) \in \mathcal{P}_a$. When $z_t^{nb} = 1$, the pump is ON and the water flows from node n to node b. When $z_t^{nb} = 0$, the pump is OFF and there is no water flow between n and b. For the pump-hosting edges, (2) can be rewritten by including the pump switching as

$$0 \leq q_t^{nb} \leq q^{\max} z_t^{nb}, \ \forall (n,b) \in \mathcal{P}_a, \ t \in \mathcal{T}. \tag{5}$$

Switching Constraints: Frequent switching of pump between ON and OFF states is not a desirable phenomenon in WDN, as it increases the transients in the network. Hence, we restrict the number of the pump switching over a predetermined time horizon (T_s). To constrain the switching, we introduce the binary toggle variables $d_t^{nb} \in \{0,1\}, \ \forall t \in \mathcal{T}, \ (n,b) \in \mathcal{P}_a$, where d_t^{nb} indicates a toggle in the state of the pump, i.e., $d_t^{nb} = 1$ when $z_t^{nb} \neq z_{t+1}^{nb}$, whereas $d_t^{nb} = 0$ implies no switching, that is the pump state is same at time indices t and $t+1$.

We observed that the switching constraints involving continuous and binary variables can be formulated with *mixed logical dynamics* [10] [11]. To this end, we introduce an auxiliary variable $\gamma_t^{nb} \in \{-1, 0, 1\}$ to model the updates of z_t^{nb} as a function of d_t^{nb}:

$$z_{t+1}^{nb} = z_t^{nb} + \gamma_t^{nb}, \tag{6}$$

where γ_t^{nb} is given by

$$\gamma_t^{nb} = \begin{cases} -1, & \text{if } d_t^{nb} = 1 \wedge z_t^{nb} = 1 \\ +1, & \text{if } d_t^{nb} = 1 \wedge z_t^{nb} = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

The logic relationship between (6) and (7) can be added to an optimization framework using the following linear inequalities [11]:

$$\mathbf{B}[\gamma_t^{nb} \ z_t^{nb} \ d_t^{nb}]^\top \leq \boldsymbol{b}, \tag{8}$$

where $\mathbf{B} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & -1 \\ 1 & 2 & 2 \\ -1 & -2 & -2 \end{bmatrix}$ and $\boldsymbol{b} = [0 \ 0 \ 3 \ 1]^\top$.

Using the toggle d_t^{nb}, the total number of switching s_t^{nb} over a time window T_s, can be written recursively as

$$s_{t+1}^{nb} = s_t^{nb} + d_t^{nb} - d_{t-T_s}^{nb}, \quad \forall t \in \mathcal{T}. \tag{9}$$

To control the number of switchings over T_s, we impose following constraint:

$$s_t^{nb} \leq \overline{s^{nb}}, \quad \forall t \in \mathcal{T}, \tag{10}$$

where $\overline{s^{nb}}$ is the maximum number of switching for the pump in the edge $(n,b) \in \mathcal{P}_a$. Finally, we assume the following initializations for the variable z_t^{nb}, v_t^{nb}, and $s_t^{nb}, \forall (n,b) \in \mathcal{P}_a$:

$$z_0^{nb} = z_{\text{init}}, \quad v_0^{nb} = v_{\text{init}}, \quad s_0^{nb} = s_{\text{init}}. \tag{11}$$

Let $\boldsymbol{z} \in \{0,1\}^{T|\mathcal{P}_a|}$ and $\boldsymbol{q} \in \mathbb{R}_+^{T|\mathcal{P}_a|}$, where $\mathbb{R}_+$ is the set of non-negative real numbers, be the vector obtained by stacking z_t^{nb} and q_t^{nb} in the lexicographical order of t, n, and b. A similar stacking of other variables $d_t^{nb}, s_t^{nb}, \gamma_t^{nb}$, and v_t^{nb} is done to obtain the vectors $\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\gamma}$, and $\boldsymbol{v}$ respectively having length $T|\mathcal{P}_a|$.

Mixed-Integer Optimization Framework: We assume that the time-ahead *dynamic energy cost* $\{\pi_t\}_{t=1}^T$ for the pump operation is given. Then, the total cost associated with pumping is given by $f_o(\boldsymbol{z}) = \sum_{t=0}^{T-1} \sum_{(n,b)\in\mathcal{P}_a} {z_t}^{nb} \pi_t$. We propose a MIO framework with the following objectives: **i)** to compute the optimal switching trajectories for the pumps that minimize $f_o(\boldsymbol{z})$ and **ii)** to ensure that while optimizing the switching strategy, all the WDN parameters are within the admissible operation range given by the equations (3),(4),(5),(6),(8),(9), (10), and (11). The proposed MIO framework is

$$\begin{aligned} \text{minimize} \quad & f_o(\boldsymbol{z}) = \sum_{t=0}^{T-1} \sum_{(n,b)\in\mathcal{P}_a} {z_t}^{nb} \pi_t \\ \text{over} \quad & \{\boldsymbol{z}, \boldsymbol{q}, \boldsymbol{v}, \boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\gamma}\} \\ \text{subject to} \quad & (3),(4),(5),(6),(8),(9),(10),(11). \end{aligned} \tag{12}$$

It is to be remarked that the switching constraint (10) controls the number of switchings over a specified time interval, which is a tedious task in manually operated WDNs.

Since the variables $\boldsymbol{z}$ and $\boldsymbol{d}$ are binary, and $\boldsymbol{\gamma}$ and $\boldsymbol{s}$ are integers, problem (12) is a MIO, which is an NP-hard nonconvex problem. Despite MIO being intractable, there are several algorithms that can be used to solve the problem approximately, among which the branch and bound algorithm and the cutting plane method are commonly used. Notably, such formulations can be solved using MIO solvers such as GLPK, Gurobi [8], etc.

III. A DATA-DRIVEN APPROACH

To this end, Section II formulates a MIO problem, where the decision variable $\boldsymbol{z}$ is computed to obtain the optimal trajectory of pump switching. Computing the decision variables of such problems is inefficient in terms of memory and computation time. In this section, we describe model-free data-driven approach to estimate the decision variables of the MIO formulation for WDN. We solve many problem instances of (12) using a standard MIO solver. The obtained solutions are used to train a feed-forward DNN which bypasses the need of a computationally expensive solver for WDN predictions.

III-A. DNN for Estimation of Decision Variables

A feed-forward DNN architecture consists of L *layers*, which define a composition of functions of the form $\hat{f}(\Theta) = h_L(h_{L-1}(\cdots h_1(\Theta))$, where Θ is the input of the DNN, $l = 1$ is the input layer, $l = 2, \cdots, L-1$ are the hidden layers and $l = L$ is the output layer. Each layer depends on the previous layer by $\boldsymbol{y_l} = \boldsymbol{h_l}(\boldsymbol{y_{l-1}}) = \sigma_l(\boldsymbol{w_l}\boldsymbol{y}_{l-1} + \boldsymbol{b}_l) \in \mathbb{R}^{N_l}$, where $\mathbb{R}$ is the set of real numbers, σ is a non-linear activation function, $\boldsymbol{w}_l$ is the weight of the neural network, and N_l is the number of nodes in layer l. The weights $\boldsymbol{w}_l$'s are obtained by training the DNN using the training data sets obtained by solving different problem instances of (12) with the help of the Gurobi MIO solver.

We use rectified linear unit (ReLU) and leaky rectified linear unit (Leaky ReLU) as activation for hidden layers. For the outer layer, sigmoid activation function is used motivated by the nature of the output decision variable and the cost is computed using binary cross entropy. Further, we apply batch normalization at the hidden layers to standardize the preactivation distribution and reduce the internal covariance shift [12]. We apply $\{v_o, z_o, s_o, d_o, \pi_t, q_t^k\}$ at the input layer of DNN for $T = \{12, 15, 18, 20, 24\}$, and the estimates of the decision variables $\{\hat{\boldsymbol{z}}_t\}$ are obtained at the output layer $l = L$ for each T.

III-B. Suboptimality

Let $f_o(\boldsymbol{z}^\star)$ denotes the optimal value of the objective function obtained by solving the MIO problem (12), whereas $f_o(\hat{\boldsymbol{z}})$ denotes the value of the objective function computed through the trained DNN model. We define the suboptimality Υ_o as

$$\Upsilon_o = \frac{|f_o(\boldsymbol{z}^\star) - f_o(\hat{\boldsymbol{z}})|}{f_o(\boldsymbol{z}^\star)} \tag{13}$$

To evaluate the performance of the DNN based approach, we consider that the estimated solution is accurate if suboptimality $\Upsilon_o \leq \epsilon$, where ϵ is the error tolerance.

IV. EXPERIMENTAL RESULTS

We use the WDN shown Fig. 2a, which consists of two reservoirs, four fixed-speed pump, two water storage tanks, and two points of water demand (users), all connected through loss-less pipes. We use the WDN parameters as: $v^{\min} = 30\,000\,\text{m}^3$, $v^{\max} = 100\,000\,\text{m}^3$, and $q^{\max} = 20\,000\,\text{m}^3/\text{h}$. We initialized using $z_{\text{init}} = 0$, $T_5 = 10$, $v_{\text{init}} = 5000\,\text{m}^3$, and $s_{\text{init}} = 3$. In addition, we considered five different time horizons $T = \{12, 15, 18, 20, 24\}$. We consider the sampling time of $\tau = 1\,\text{h}$, $\forall\ T$. The electric energy cost π_t, $\forall t \in \mathcal{T}$ and the demand

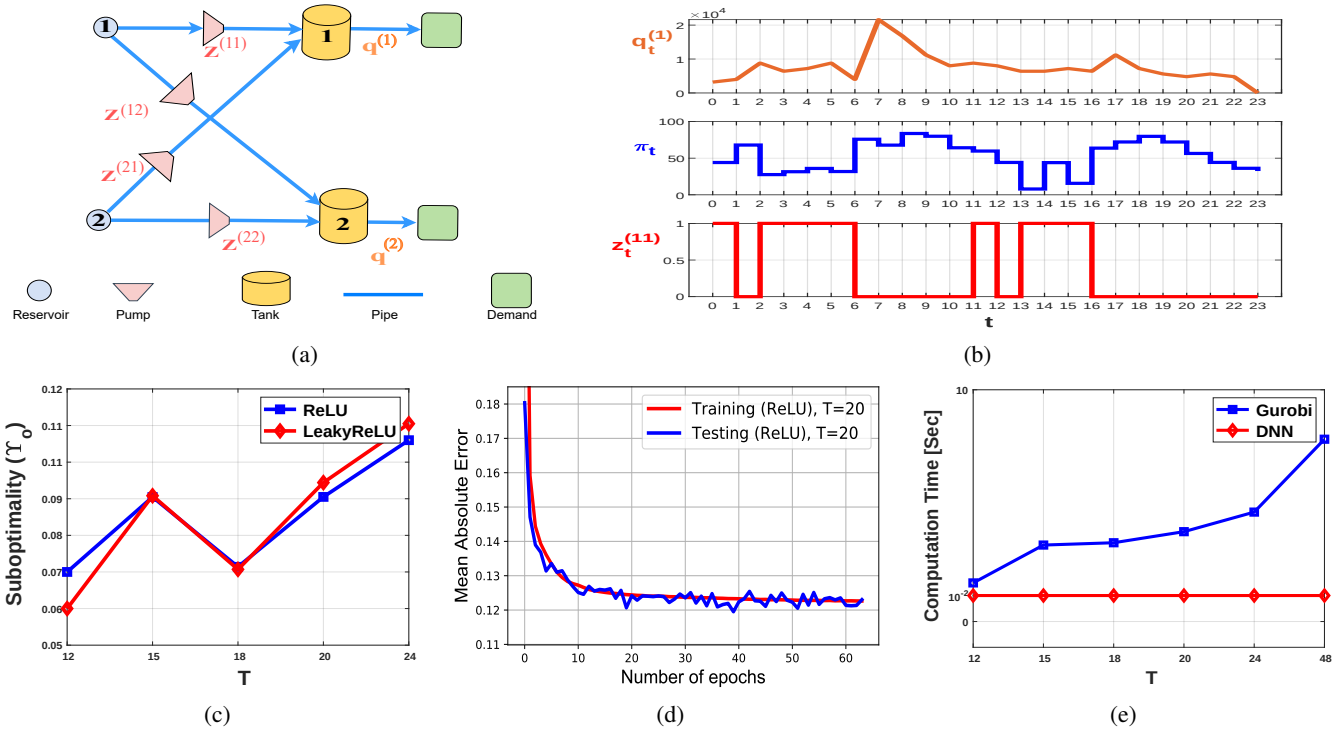


Fig. 2: (2a.) depicts the schematic representation of the WDN, (2b.) presents the pump scheduling z_t and other WDN variables for $T = 24$ with $\tau = 1$hr, (2c.) presents the suboptimality comparison for ReLU and LeakyReLU, (2d.) shows the convergence of DNN for activation ReLU at T=20, and (2e.) presents the comparison of computation time between DNN and Gurobi.

profile of user consumption at the point of water demand q_t^{bn}, $\forall t \in \mathcal{T}$, $(b, n) \in \mathcal{P} \setminus \mathcal{P}_a$ are generated synthetically by assuming a daily pattern of user consumption and electric energy cost for 200000 problem instances for each time horizons T. Also, we added the switching constraints $\overline{s^{nb}} = 5$ and $T_s = 10$, meaning that total number of switching over a time window of length 10 should not exceed 5.

The MIO problem (12) is solved using the Gurobi solver to obtain the pump trajectory $z^\star$. Fig. 2b depicts the obtained pump switching trajectory of a problem instance for a time window of 24 hours. As expected, the pump $z_t^{(11)}$ is in ON state when the energy cost π_t is smaller. The data, which contains all the above mentioned WDN variables and the Gurobi solutions are split as $90\% - 10\%$ for training and testing the DNN. We train the DNN using random initialization with ADAM optimizer. The structure of DNN is constructed by one input layer, 8 hidden layers and one output layer. The hidden layers were consisting of 20, 40, 60, 80, 100, 80, 60, and 40 neurons respectively. In addition, we fixed the value of the Leaky ReLU parameter α as 0.1.

The model performance is evaluated using the suboptimality Υ_o defined in (13) averaged over the test data set and the results given in Fig. 2c shows that the feed-forward DNN can estimate the decision variables of MIO formulation within a tolerance range of $\epsilon \leq 11 \times 10^{-2}$. Fig. 2d shows the convergence of the DNN cost function for $T = 20$. Fig. 2e compares the computation time of the Gurobi solver and the proposed DNN-based solver. Table I presents the percentage of instances for which constraint (5) is satisfied for $\{\hat{z}_t\}$, which ranges between 85% to 90%. We would like to leave a remark that the constraint violation can be addressed by projecting the affected variables of the corresponding problem instances to the feasible region defined by the constraints. The tests were conducted on a 2.7 GHz, Intel Core i7 computer with 8 GB RAM. It is observed that there is an exponential increase in computation time for Gurobi solver when $T \geq 48$, whereas the DNN solves the problem in milliseconds for all values of T.

Table I: Constraint Satisfaction

T	*ReLU*	*LeakyReLU*
12	85.57 %	85.12 %
15	89.5 %	89.28 %
18	81.87 %	81.5 %
20	87.8 %	87.74 %
24	89.5 %	89.88 %

V. CONCLUSION

With ever increasing expansion of WDN, pump scheduling using existing MIO solver is inefficient in terms of memory and computational time. In this work, we propose a MIO formulation to minimize the electric energy cost of WDN while keeping the WDN parameters within a desired admissible range. Further, we propose a computationally efficient method to solve the MIO formulation using DNN, which can bypass the need of using MIO solvers. In a real WDN, given the various interconnected tanks, valves, pump, user demand and dynamic energy price, this approach could compute the decision variables in a computationally efficient and feasible manner, given the availability of necessary data.

VI. REFERENCES

[1] M Fayzul K Pasha and Kevin Lansey, "Strategies to develop warm solutions for real-time pump scheduling for water distribution systems," *Water resources management*, vol. 28, no. 12, pp. 3975–3987, 2014.

[2] Paul F Boulos, Zheng Wu, Chun Hou Orr, Michael Moore, Paul Hsiung, and Devan Thomas, "Optimal pump operation of water distribution systems using genetic algorithms," in *Distribution system symposium*. Citeseer, 2001.

[3] Shen Wang, Ahmad Taha, Nikolaos Gatsis, and Marcio Giacomoni, "Receding horizon control for drinking water networks: The case for geometric programming," *IEEE Transactions on Control of Network Systems*, 2020.

[4] Dariush Fooladivanda and Joshua A Taylor, "Energy-optimal pump scheduling and water flow," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1016–1026, 2017.

[5] Manish K Singh and Vassilis Kekatos, "Optimal scheduling of water distribution systems," *IEEE Transactions on Control of Network Systems*, 2019.

[6] Dimitris Bertsimas and Robert Weismantel, *Optimization over integers*, vol. 13, Dynamic Ideas Belmont, 2005.

[7] Dimitris Bertsimas and Bartolomeo Stellato, "Online mixed-integer optimization in milliseconds," *arXiv:1907.02206*, 2019.

[8] LLC Gurobi Optimization, "Gurobi optimizer reference manual," 2020.

[9] Kevin E Lansey and LW Mays, "Optimal design of water distribution systems," *Water Distribution System Handbook, McGraw-Hill, New York*, 2000.

[10] Alberto Bemporad and Manfred Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, 1999.

[11] Damian Frick, Alexander Domahidi, and Manfred Morari, "Embedded optimization for mixed logical dynamical systems," *Computers & Chemical Engineering*, vol. 72, pp. 21–33, 2015.

[12] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*. PMLR, 2015, pp. 448–456.

Cooperative Communication, Localization, Sensing and Control for Autonomous Robotic Networks

Siwei Zhang, Emanuel Staudinger, Robert Pöhlmann and Armin Dammann
Institute of Communications and Navigation, German Aerospace Center (DLR), Oberpfaffenhofen, Germany
Email: firstname.lastname@DLR.de

***Abstract*—Networks composed of a myriad of autonomous robots have attracted increasing attention in recent years due to their enormous capability expansion from single robot systems. In these networks, robots benefit from the collaboration with each other to enhance their situation awareness for autonomous operation. For example, in an extraterrestrial exploration mission, a robotic swarm can collaboratively utilize the inter-robot communication system to propagate information, synchronize itself, and navigate to achieve mission objectives like joint environmental sensing. In addition, each robot can decide and control its own trajectory, so that the aforementioned tasks are accomplished in a globally efficient manner. In this paper, we propose multi-agent control strategies for autonomous robotic networks, which adapt the mission demands on cooperative communication, localization and sensing. We also discuss three space exploration examples with different mission demands, which lead to distinct network formations. These three missions will be conceptually demonstrated in a space analog mission on the volcano Mount Etna in June 2022.**

I. Introduction

A swarm of autonomous robots [1], analog to a biological swarm in nature [2], can rapidly explore a vast area on earth or in space, make simultaneous observations at different locations, and avoid a single point of failure. Therefore, it is a promising concept and a paradigm shift in exploration of human inaccessible area, e.g. search-and-rescue [3], environmental monitoring [4] and future space missions [1], [5]. A major challenge of such an exploration is communication and navigation, since external infrastructures like cellular networks or global navigation satellite systems (GNSSs) are often absent in a human inaccessible area. In this case, the network composed of a myriad of robots can additionally provide communication and localization service in the area as a temporal infrastructure, next to its exploration tasks like collaborative environment sensing and target tracking. Regarding autonomy, every robot is free to choose its trajectory, jointly considering the requirements of cooperative communication, localization and exploration. It is a multi-agent, multi-objective optimization problem, which often needs to be solved in-situ due to the unpredictable environmental situations. In [6] we have proposed an autonomous swarm navigation framework, where a swarm optimizes its formation to improve localization, while achieving mission objectives.

In this paper, we have a joint look at the communication, localization, sensing and control aspects in an exploration mission, which is essential for design such an autonomous robotic network [7], [8]. Depending on the overall mission objectives, the trade-off between communication and localization determines the favorable swarm formations. We investigate this trade-off with three representative swarm exploration missions, namely remote rover manipulation, geological exploration and low frequency array (LOFAR). We extend the autonomous swarm navigation framework from [6] with cooperative communication. It allows the swarm to adapt its formation according to the communication-localization trade-off, ergo the mission objectives. Both simulations and experiments have been conducted to prove the concept of autonomous swarm navigation, which will be demonstrated in the space-analog mission planned on volcano Mount Etna, Italy, in 2022 [9].

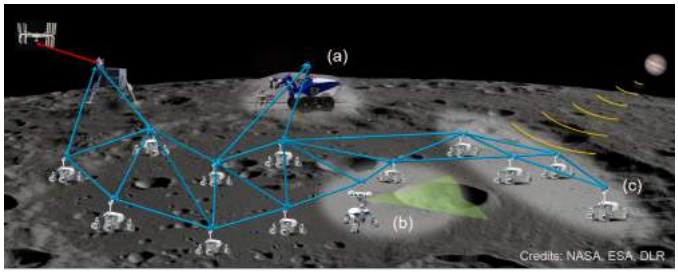


Figure 1: Concepts of three lunar swarm missions: (a) remote rover manipulation, (b) geological exploration and (c) LOFAR.

II. Lunar Swarm Exploration Concepts

We study three lunar swarm exploration concepts, as shown in Figure 1. A swarm of rovers move from the mission base to operation areas, while maintaining localizability and communication.

In the first mission, one rover at area (a) is steered remotely by astronauts from an orbiter. For this mission, the localization requirement is relatively low. However, a reliable communication between the rover and the orbiter needs to be guaranteed for real-time video and sensor data streaming.

In the second mission at area (b), two rovers are mapping the environment for geological exploration. In order to collectively reconstruct the environmental map, large volume of sensor data, together with its precise collecting location, need to be frequently exchanged. In this mission, both communication and localization are demanding.

In the third mission, four rovers at area (c) form a LOFAR for radio astronomy. As an example, they receive the radio bursts of Jupiter and determine the direction of Jupiter with respect to (w.r.t.) the swarm. In this case, the location

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

estimates of the rovers have to be sufficiently precise, while communication is less demanding.

In these lunar exploration concepts, the communication-localization trade-off varies according to the application. Next, we look into the design of the swarm control, optimizing the swarm formation for cooperative localization, communication and mission objectives like target tracking, respectively.

III. Multi-Agent Control

Let us consider a swarm of robots with their positions at time t collectively denoted as $\mathbf{p}_{\mathcal{A}}^{(t)}$. Robots' positions can be manipulated with a control command $\mathbf{u}_{\mathcal{A}}^{(t)}$ into $\mathbf{p}_{\mathcal{A}}^{(t+1)}$ as

$$\mathbf{p}_{\mathcal{A}}^{(t+1)} = f(\mathbf{p}_{\mathcal{A}}^{(t)}, \mathbf{u}_{\mathcal{A}}^{(t)}, \boldsymbol{\epsilon}_{\mathcal{A}}^{(t)}), \tag{1}$$

where $\boldsymbol{\epsilon}_{\mathcal{A}}^{(t)}$ denotes the controller noise. As an example of sensing applications, we consider tracking a target at position $\mathbf{p}_{\mathrm{t}}^{(t)}$. A robot in the network is referred to as 'leader' with position $\mathbf{p}_{\mathrm{l}}^{(t)}$, which aims at following the target. Another static robot at position $\mathbf{p}_{\mathrm{s}}^{(t)}$ close to the mission base is referred to as 'sink'. As an example of communication requirements, we consider nonspecific sensor data about the target transmitted from the leader to the sink over the network. Communication and distance measurements between entities can be conducted if the entities are within their respective communication range. Three static entities close to the mission base serve as anchors with known positions. Both positions of the robots and targets need to be estimated with distance measurements in the network $\mathbf{z}^{(t)}$. The overall executed control command $\tilde{\mathbf{u}}_{\mathcal{A}}^{(t)}$ can be derived, for example with a linear combination of the controls for localization $\mathbf{u}_{\mathrm{L}}^{(t)}$, target tracking $\mathbf{u}_{\mathrm{T}}^{(t)}$ and communication $\mathbf{u}_{\mathrm{C}}^{(t)}$. The individual control commands are introduced next.

A. Cooperative Localization

For cooperative localization, we aim at finding an optimized control command $\mathbf{u}_{\mathrm{P}}^{(t)}$, so that the predicted measurements $\tilde{\mathbf{z}}^{(t+1)}$ provide the richest information for estimating the new positions $\mathbf{p}_{\mathcal{A}}^{(t+1)}$, i.e.

$$\mathbf{u}_{\mathrm{P}}^{(t)} = \underset{\mathbf{u}_{\mathcal{A}}^{(t)}}{\arg\min}\ \mathbb{E}\Big[\|\mathbf{p}_{\mathcal{A}}^{(t+1)} - \hat{\mathbf{p}}_{\mathcal{A}}^{(t+1)}\|^2\Big]. \tag{2}$$

We employ the information seeking control proposed in [6] to solve (2). The estimation uncertainty is inferred with the Cramér-Rao bound (CRB), i.e.

$$\mathrm{CRB}\Big[\mathbf{p}_{\mathcal{A}}^{(t+1)}\Big] \preccurlyeq \mathbb{E}\Big[\|\mathbf{p}_{\mathcal{A}}^{(t+1)} - \hat{\mathbf{p}}_{\mathcal{A}}^{(t+1)}\|^2\Big]. \tag{3}$$

The optimized control command is proportional to the negative gradient of the CRB, i.e.

$$\mathbf{u}_{\mathrm{L}}^{(t)} \propto -\nabla_{\mathbf{u}_{\mathcal{A}}^{(t)}} \mathrm{CRB}\Big[\mathbf{p}_{\mathcal{A}}^{(t+1)}\Big]. \tag{4}$$

B. Cooperative Target Tracking

The objective of cooperative target tracking is twofold. Firstly, all the robots compose a sensor array whose formation is favorable for estimating the position of the target, hence minimizing the position CRB of the target:

$$\mathbf{u}_{\mathrm{T}}^{(t)} \propto -\nabla_{\mathbf{u}_{\mathcal{A}}^{(t)}} \mathrm{CRB}\Big[\mathbf{p}_{\mathrm{t}}^{(t+1)}\Big]. \tag{5}$$

Secondly, the leader moves towards the target with an additional control command

$$\mathbf{u}_{\mathrm{L}}^{(t)} = \underset{\mathbf{u}_{\mathcal{A}}^{(t)}}{\arg\min} \|\mathbf{p}_{\mathrm{l}}^{(t+1)} - \mathbf{p}_{\mathrm{t}}^{(t+1)}\|. \tag{6}$$

C. Cooperative Communication

The robotic network can be considered as an undirected graph with robots as the vertices. If two vertices i and j are within the communication range of each other, they are connected with an edge l_{ij}. The maximum throughput on this edge is expressed with the Shannon capacity:

$$C_{ij} = B \log_2(1 + \mathrm{SNR}_{ij}), \tag{7}$$

where B is the bandwidth and SNR_{ij} is the signal to noise ratio (SNR) of the edge which is proportional to the Euclidean distance d_{ij} between vertices i and j. We assume a decode-and-forward relaying scheme for all the robots. The control strategy for communication can be formulated as finding a suitable route $\mathcal{L} = \{\cdots l_{ij}, \cdots\}$ from the leader to the sink, and determining a control command $\mathbf{u}_{\mathrm{C}}^{(t)}$ that minimizes all edge distances in the route, i.e.

$$\mathbf{u}_{\mathrm{C}}^{(t)} = \underset{\mathbf{u}_{\mathcal{A}}^{(t)}}{\arg\min}\{d_{ij} :\ \forall l_{ij} \in \mathcal{L}\}. \tag{8}$$

A suitable route can either be a single path connecting the leader and the sink, or include multiple paths. Multiple strategies can be exploited for finding a suitable route.

1) Pre-Defined Path: A naive strategy is to pre-define a fixed route from the leader to the sink. For example, half of the robots in the network are chosen as relays.

2) Shortest Path: The shortest path is the one with minimal number of relays in between. The Dijkstra's algorithm [10] can be applied to find the shortest path. This strategy is suitable when the communication requirement is less demanding like in the LOFAR mission.

3) Widest Path: The widest path is a single path which has the highest throughput. Therefore it is also referred to as the maximum capacity problem [11]. For the decode-and-forward relaying model, the total throughput equals the minimum edge capacity. Hence finding the widest path is equivalent to finding a path whose longest edge is the shortest among all possible paths. It can be achieved with a modification of the Dijkstra's algorithm. A route with widest path is suitable when the communication requirement is as demanding as other requirements like in the geological exploration mission.

4) Maximum Flow: The maximum flow strategy exploits the whole network for communication with high priority. It maximizes the total throughput allowing multiple paths. The Boykov-Kolmogorov algorithm [12] can be utilized to calculate the route with maximum flow. This strategy is suitable for

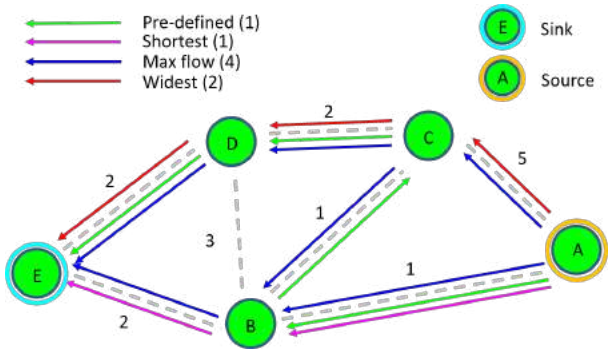


Figure 2: Different routing strategies for a toy scenario.

demanding communication requirements like the rover remote manipulation mission.

The above-mentioned routing strategies are illustrated in Figure 2. The edge capacity is labelled at the individual edge. The total route capacities are shown in the legend.

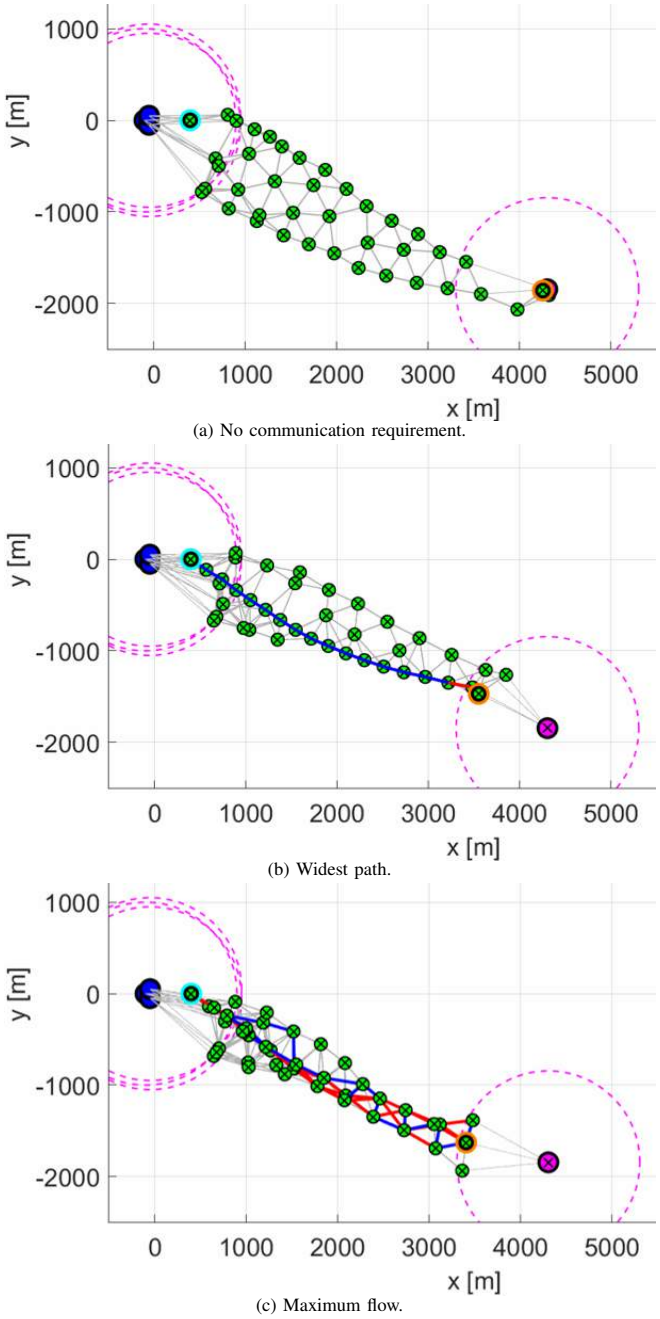
IV. Results

A. Simulations

We conduct swarm control simulations taking five different strategies for cooperative communication, namely no communication requirement considered (shortest path for calculating throughput), pre-defined path, shortest path, widest path and maximum flow. Three simulation snapshots are shown in Figure 3. A swarm of 40 robots (green markers) depart from three anchors (blue markers) to track a target (magenta markers) beyond the communication range of the anchors (magenta dashed lines). The leader and sink are labelled with orange and cyan circles, respectively. Edges with distance measurements, communication and communication with full capacity are indicated with gray, blue and red lines, respectively. In all three scenarios, robots build a rigid bridge connecting anchors, swarm and the target, so that the swarm and target are localizable. In the case of no communication requirement, the leader reaches the target. In the case of widest path, the relays form almost a straight line to have a high total path capacity. In the case of maximum flow, the route contains multiple paths with high relay concentration. In this way the total route capacity is maximized. A four dimensional performance comparison of different swarm control strategies, normalized to the respective maximum value, is shown in Figure 4. Maximum flow based control guarantees the highest throughput at the cost of having the largest number of hops. No communication requirement or shortest path based control experience a large communication latency, but consume less hops and are more efficient in localization and target following. Widest path based control is well balanced in all four metrics and outperforms the pre-defined path approach. As a conclusion, the proposed strategies can be used in swarm control, according to the desired communication-localization trade-off.

B. Experiments

Within the project Autonomous Robotic Networks to Help Modern Societies (ARCHES) we will demonstrate robotic



(a) No communication requirement.

(b) Widest path.

(c) Maximum flow.

Figure 3: Swarm formation snapshots with different cooperative communication strategies.

exploration technologies in a lunar analog environment on Mount Etna, Italy, in 2022 [9]. The demonstration mission consists of three scenarios that are closely related to the three conceptual missions introduced in Section II. The first two scenarios examine technical and operational aspects of geological in-situ analysis and sample return, with rovers either manipulated remotely by an ESA astronaut, or operating autonomously. In these scenarios, static boxes or rovers are deployed at preferable positions as relays to guarantee reliable communication. The third scenario demonstrates the autonomous deployment of a LOFAR, which is illustrated in

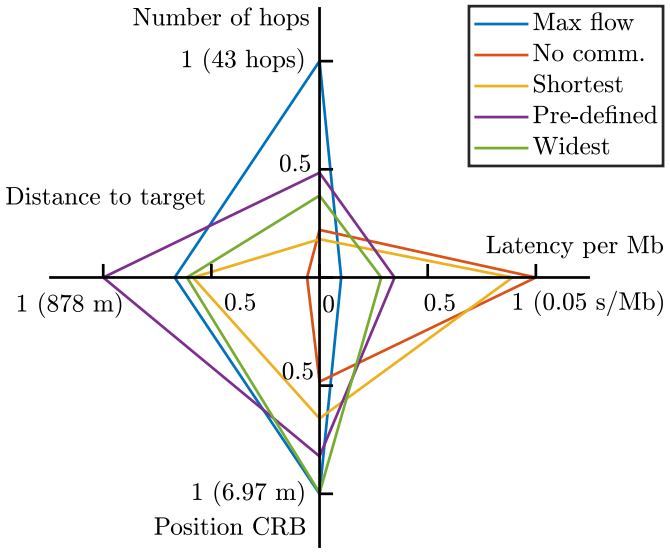


Figure 4: Performance comparison of different swarm control strategies for communication.

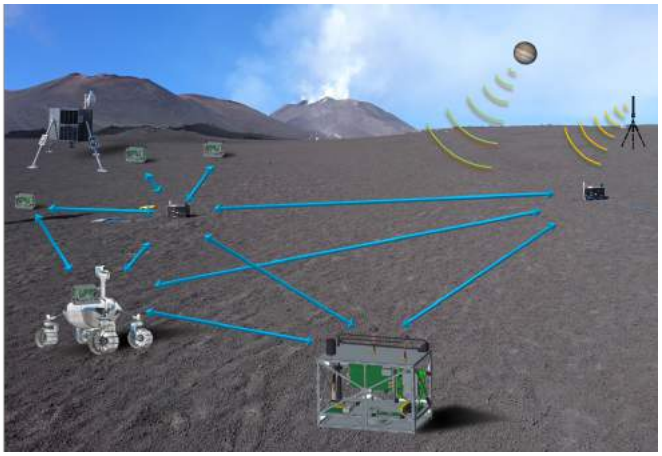


Figure 5: Space-analog LOFAR mission in 2022.

Figure 5. LOFAR payload boxes are placed by a lightweight rover and precisely synchronized and localized with our swarm navigation system [8]. Low frequency radio signals emitted either from space or by an artificial transmitter are detected by this array. As we discussed in Section II, in this scenario the most demanding requirement is localization.

As a preparation to this space analog mission, we have conducted a LOFAR experiment with our developed swarm navigation platform [7] at the German Aerospace Center (DLR) in March 2021. The experimental setup with four LOFAR boxes and two low frequency transmitters is shown in Figure 6. Software defined radio is used to generate and receive radio signals. Among LOFAR boxes and anchors, radio signals with a carrier frequency around 5.5 GHz and a bandwidth of 25 MHz are transmitted for time of flight (ToF) based distance measurement. From the low frequency transmitters, sine-waves with a carrier frequency around 20 MHz are transmitted. A real-time decentralized particle filter [13] is implemented for each LOFAR box for swarm cooperative localization. The

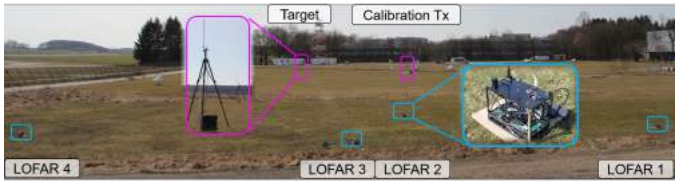


Figure 6: Swarm self localization experiment as preparation of LOFAR mission.

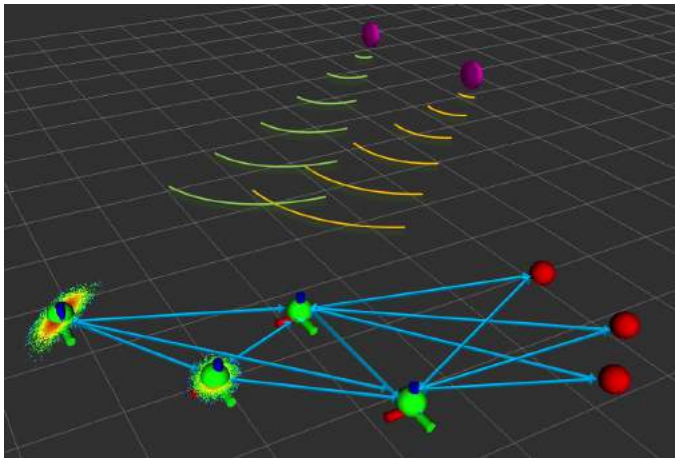


Figure 7: Experimental results for swarm localization.

localization result can be seen in Figure 7. The ground-truth positions of three anchors, four LOFAR boxes and two low frequency transmitters are measured with the GNSS-real-time kinematic (RTK) systems, illustrated with red markers, axes, and magenta markers. The point estimates of the LOFAR boxes' positions are indicated with green markers, surrounded by particles. Due to the curvature of the ground and low box's elevation, LOFAR boxes three and four cannot obtain distance measurements from any anchors. However, they are still able to estimate their positions through cooperation with LOFAR boxes one and two, with larger position uncertainty. Sub-meter position uncertainty can be obtained even for LOFAR boxes that do not directly connected with anchors. This accuracy is at least an order of magnitude smaller than the wavelength of the low frequency signals (15 m), ergo sufficient for LOFAR mission [8]. The LOFAR signals are detected by all LOFAR boxes coherently. As an ongoing work, we are analysing the performance of the direction of arrival (DoA) estimation of the low frequency transmitters.

V. Conclusion

In this paper, we proposed a swarm control framework which jointly considers the mission requirements for communication, localization and sensing. Multiple strategies can be chosen according to the desired communication-localization trade-off for different space exploration missions. Simulations verify the proposed swarm control framework, while experiments prove the concept of swarm cooperative localization in a LOFAR mission which will be demonstrated in a space analog mission on Mount Etna in 2022.

ACKNOWLEDGMENT

Part of the presented research has been supported by the Helmholtz Association project ARCHES (contract number ZT-0033). The authors would like to thank Kimon Cokona for his support.

REFERENCES

[1] E. Vassev, R. Sterritt, C. Rouff, and M. Hinchey, "Swarm technology at NASA: Building resilient systems," *IEEE IT Prof.*, vol. 14, no. 2, pp. 36–42, Mar. 2012.

[2] M. Ballerini *et al.*, "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

[3] M. Bernard, K. Kondak, I. Maza, and A. Ollero, "Autonomous transportation and deployment with aerial robots for search and rescue missions," *J. Field Robot.*, vol. 28, no. 6, pp. 914–931, 2011.

[4] M. Dunbabin and L. Marques, "Robots for environmental monitoring: Significant advancements and applications," *IEEE Robot. Autom. Mag.*, vol. 19, no. 1, pp. 24–39, Mar. 2012.

[5] A. Seeni, B. Schfer, and G. Hirzinger, "Robot mobility systems for planetary surface exploration – state-of-the-art and future outlook: A literature survey," in *Aerospace Technologies Advancements*, T. T., Ed. London: InTech, Jan. 2010, pp. 189–208.

[6] S. Zhang, R. Pöhlmann, T. Wiedemann, A. Dammann, H. Wymeersch, and P. A. Hoeher, "Self-aware swarm navigation in autonomous exploration missions," *Proc. IEEE*, vol. 108, no. 7, pp. 1168–1195, 2020.

[7] S. Zhang, R. Pöhlmann, E. Staudinger, and A. Dammann, "Assembling a swarm navigation system: Communication, localization, sensing and control," in *2021 IEEE 18th Annual Consumer Communications Networking Conference (CCNC)*, 2021, pp. 1–9.

[8] E. Staudinger, S. Zhang, R. Poehlmann, and A. Dammann, "The role of time in a robotic swarm: A joint view on communications, localization, and sensing," *IEEE Commun. Mag.*, vol. 59, no. 2, pp. 98–104, 2021.

[9] M. J. Schuster *et al.*, "The ARCHES space-analogue demonstration mission: Towards heterogeneous teams of autonomous robots for collaborative scientific sampling in planetary exploration," *IEEE Robotics and Automation Letters*, pp. 1–1, 2020.

[10] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, no. 1, p. 269–271, Dec. 1959. [Online]. Available: https://doi.org/10.1007/BF01386390

[11] M. Pollack, "The maximum capacity through a network," *Operations Research*, vol. 8, no. 5, pp. 733–736, 1960. [Online]. Available: http://www.jstor.org/stable/167387

[12] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.

[13] S. Zhang, E. Staudinger, T. Jost, W. Wang, C. Gentner, A. Dammann, H. Wymeersch, and P. A. Hoeher, "Distributed direct localization suitable for dense networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1209–1227, 2020.

FIRST STEPS TOWARD THE DEVELOPMENT OF VIRTUAL PLATFORM FOR VALIDATION OF AUTONOMOUS WHEEL LOADER AT PULP-AND-PAPER MILL: MODELLING, CONTROL AND REAL-TIME SIMULATION

Michael A. Kerr[1]*, Danielle S. Nasrallah*[2,*]*, and Tsz-Ho Kwok*[1]

[1]Department of Mechanical, Industrial and Aerospace Engineering, Concordia University, QC, Canada
[2]Advanced Control and Intelligent Mobility, OPAL-RT Technologies, QC, Canada

ABSTRACT

The forestry industry all over the world is seeing the need for modernization of its machines toward autonomy. In this paper, we focus on a wheel loader that should operate autonomously in the yard of a pulp-and-paper mill, scooping wood chips from a pile of wood and dropping them into a hopper, which is linked to a conveyor that carries them inside the mill. The modelling of the wheel loader is elaborated first, while taking into account that it is composed of two systems: (i) the vehicle and (ii) the arm carrying the bucket. Notice that the former pertains to the category of articulated vehicles that steer using a different mechanism than the conventional Ackermann steering used in car-like vehicles. As for the latter, it is a 2DOF serial manipulator. The navigation is considered then. Finally, simulation results of the kinematics model are shown in Matlab/Simulink first, then dynamics and 3D animation are added using ROS2/Gazebo. Notice that this work is a first step toward the development of the digital twin of the wheel loader. Later, it will be used as the virtual platform for the validation of the autonomous wheel loader.

Index Terms— Autonomous, Off-road, Articulated Vehicles, Modelling, Control, Navigation, Real-time Simulation, Co-simulation, Digital Twin, Virtual platform, RT-LAB, ROS2, Gazebo.

1. INTRODUCTION

The modernization of forestry machines toward autonomy is becoming a MUST due to (i) the aging of the workforce, (ii) the discomfort of performing work tasks, etc. Canada, Northern Europe, Australia, Brazil, and other countries are among the major players. In this paper the application requests that a wheel loader navigates autonomously in the yard of a pulp-and-paper mill to scoop wood chips from a pile of wood and drop them into a hopper. This is a typical operation that can be visualized here for instance: https://www.youtube.com/watch?v=txBdy0m5H44

At the same time, real-time simulation (RTS) has gained acceptance as an essential tool in R&D activities over the last two decades. RTS offers many attractive features: it reduces time-to-market, allows design flexibility, and validates the distinct phases of the concept. The strength of RTS becomes apparent when the application is under development and requires an interaction with the real physical subsystems, which is the case here because the wheel loader exists already, while a new controller to make it autonomous is under development.

Notice that a similar work was undertaken by VOLVO CE Sweden in 2008 in association with many Swedish institutions. It focused on rendering autonomous a wheel loader scooping gravel and loading a hauler. This project lasted for four years and generated around 11 graduate theses. We found highly interesting material in five of them [1, 2, 3, 4, 5]. In addition, we noticed that most of the development consisted of minor simulation phases and major prototyping phases. Furthermore, after contacting directly the chief engineer of this project at Volvo, we understood that the project remained at the prototyping level.

Our vision to approach the autonomy is different as we want to develop a fully virtual simulation platform including the vehicle within its environment of operation. Then, we will go through the V-Model used for software design/development and testing/validation phases, namely, (i) Model-In-the-Loop (MIL), (ii) Software-In-the-Loop (SIL), (iii) Hardware-In-the-loop (HIL), and (iv) Vehicle-In-the-Loop (VIL). Therefore, rushing into prototyping is not our priority as we want to build a virtual platform, which is flexible enough to extrapolate beyond this specific application and generalize the usage of digital twins as an essential step in the autonomy journey. Therefore, at this early design stage, it is essential to start with a fully virtual system composed of (i) the wheel loader, (ii) controller, and (iii) the operating environment.

Having said that, the modelling of the wheel loader is achieved in Section 2 while noticing that it is an articulated vehicle that is not relying on Ackermann steering used in car-like vehicle. It was referred to as a *Laud-Haul-Dump* vehicle [6]. Then, the wheel loader arm is modelled as a 2-DOF (degree-of-freedom) robotic serial manipulator. Section

*Corresponding author. Email: danielle.nasrallah@opal-rt.com

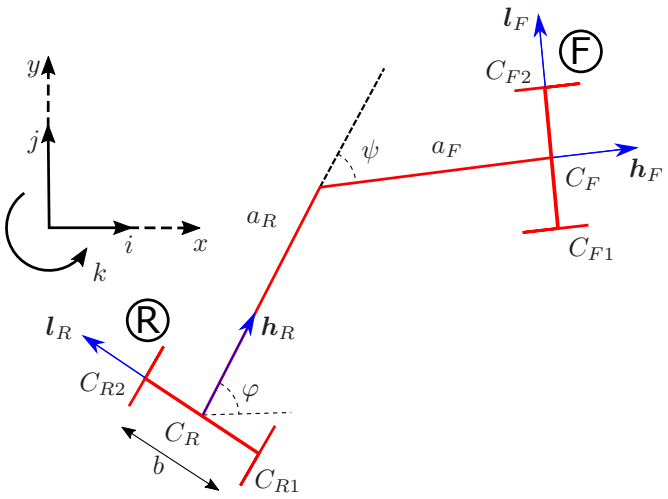


Fig. 1. Articulated vehicle diagram

3 covers navigation where similar work was done by Hitachi Construction Machinery and Japanese institutions [7], namely the v-shape loading method which is replaced by a $\gamma-$shape in our case. Simulation results using Matlab/Simulink are shown in Section 4 and the model representation in Gazebo with ROS2 is shown in 5. Section 6 concludes the paper and explain the future developments.

2. MODELLING

As stated above the wheel loader is composed of (i) the articulated vehicle and (ii) the loader arm. The vehicle, shown in Fig. 1, is composed of front and rear chassis, each with two wheels that remain parallel to the corresponding chassis. The vehicle steering is achieved by means of the revolute joint linking the two chassis. The loader arm is composed of two links with two revolute joints, namely, lift and tilt.

2.1. Articulated Vehicle

In order to elaborate the kinematics modelling of the articulated vehicle we will use the following assumptions:

1. Motion on a horizontal plane: Knowing that the operation is taking place on the yard of pulp-and-paper mill, the terrain is indeed horizontal, otherwise it becomes problematic to maintain the pile of wood chips;
2. Pure-rolling conditions: The yard is shoveled in winter time and snow salt is added to increase the wheels' adhesion to the ground; and,
3. Rear-driven vehicle: The wheel loader is a rear-driven vehicle as the mechanism of the bucket's manipulator is held by the front chassis.

Thus, the vector of generalized coordinates becomes

$$\boldsymbol{q} = \begin{bmatrix} x & y & \psi & \varphi & \theta \end{bmatrix}^T \tag{1}$$

where x and y represent the Cartesian coordinates of C_R – the midpoint of the rear wheels, in the inertial frame $\mathcal{F}_0$ represented by the right-handed orthogonal triad $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. The steering angle ψ describes the rotation of the front chassis with respect to the rear one. The angle φ represents, in turn, the orientation of the rear chassis in $\mathcal{F}_0$. Finally, the angle θ is the average angular displacement of the two rear wheels, associated directly to v – the heading speed of the vehicle. The parameters: a_R and a_F represent the length of the rear and front chassis, respectively, while b is the distance between the rear wheels. Moreover, the right-handed orthogonal triads $\{\mathbf{h}_R, \mathbf{l}_R, \mathbf{k}\}$ and $\{\mathbf{h}_F, \mathbf{l}_F, \mathbf{k}\}$ are associated to the rear and front chassis, thus describing their orientation in $\mathcal{F}_0$.

Under the pure-rolling assumptions, deriving kinematics equations leads to:

$$\dot{\varphi} = \frac{v \sin\psi - a_F \dot{\psi}}{a_R \cos\psi + a_F} \tag{2}$$

where $\dot{\psi}$ and v represent the two inputs of the vehicle, namely, the steering rate and the heading speed, respectively. That being said, the state space representation of the articulated vehicle is given by:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi} \\ \dot{\varphi} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\varphi & 0 \\ \sin\varphi & 0 \\ 0 & 1 \\ \dfrac{\sin\psi}{a_R\cos\psi + a_F} & \dfrac{-a_F}{a_R\cos\psi + a_F} \\ \dfrac{1}{r_w} & 0 \end{bmatrix} \begin{bmatrix} v \\ \dot{\psi} \end{bmatrix} \tag{3}$$

2.2. Loader Arm

As stated above, the loader arm is an open simple kinematic chain composed of two links and the base, referred to as link 0. Using the classical Denavit-Hartenberg (DH) notation [8], the frame $\mathcal{F}_1$ and its associated orthogonal triad $\{\mathbf{i}_1, \mathbf{j}_1, \mathbf{k}_1\}$ is to be attached to the base link 0. However, knowing that the base is not fixed in this specific case, as it is rigidly attached to the front chassis, one can easily notice that $\mathbf{i}_1 = \mathbf{k}$, $\mathbf{j}_1 = \mathbf{h}_F$, and $\mathbf{k}_1 = \mathbf{l}_F$. Figure 2 shows the wheel loader arm in the vertical plane $(\mathbf{k}, \mathbf{h}_F)$. Moreover, on the DH notation, the lift joint is along $\mathbf{k}_1$ while the tilt joint is along $\mathbf{k}_2$. Furthermore, O_1 denotes the origin of $\mathcal{F}_1$ while P denotes the origin of $\mathcal{F}_3$, which turns out to be also the end effector (EE).

Notice that Fig. 2 contains the additional symbols:

- φ_l, the lift angle, with: $\varphi_l = \dfrac{\pi}{2} - \theta_1$
- φ_t, the tilt angle, with: $\varphi_t = \theta_2$
- α and β angles as well as r length, that will be used later in the inverse kinematics computation.

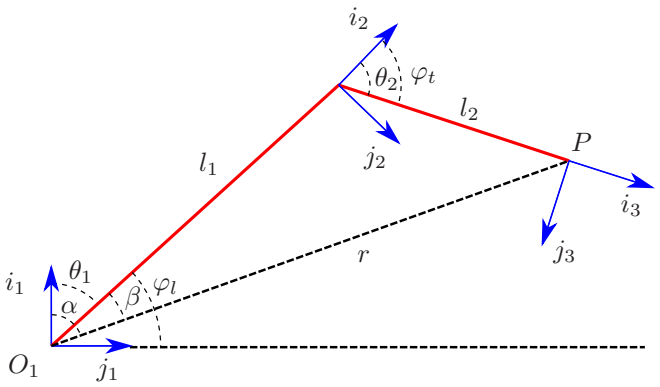


Fig. 2. Wheel loader arm, a serial two-axis manipulator

The DH-Table of the loader arm is given below:

i	a_i	b_i	α_i	θ_i
1	l_1	0	0	θ_1
2	l_2	0	0	θ_2

3. NAVIGATION

As mentioned previously, the wheel loader is supposed to move back and forth between the pile of wood chips and the hopper. In [7] a v-shape loading method was suggested. In this application we found that a γ-shape is more appropriate as shown in Fig. 3.

Fig. 3. Paper Mill Gamma Shape

Assuming the wheel loader starts between the pile of wood chips and the hopper, the wheel loader will perform the following actions

1. Forward and turn toward the wood pile with bucket in rest position
2. Approach pile and lower bucket to ground
3. Forward into pile while scooping bucket upward collecting wood chips
4. Reverse from pile and lower bucket back to rest position while not spilling wood chips
5. Reverse and turn back to start position with bucket in rest position
6. Forward and turn in opposite direction toward hopper
7. Approach hopper and raise bucket simultaneously while not spilling wood chips
8. Perform small adjustments to dump wood chips appropriately in hopper
9. Reverse and return bucket to rest position

These actions will repeat with small adjustments for (i) wood chip pile size and location of optimal scoop, (ii) amount and location of wood chips in the hopper, or (iii) other environmental factors. A simulation of these actions was modelled and animated using Matlab/Simulink with multiple scenarios:

1. pile of woods: approaching the center, slightly to the left, to the right
2. hopper: empty, quarter, half, three quarters and full

4. MATLAB/SIMULINK SIMULATION RESULTS

In this section we show the simulation results obtained in Matlab/Simulink. A top view of the terrain is given in Fig. 4 highlighting the γ-shape adopted here.

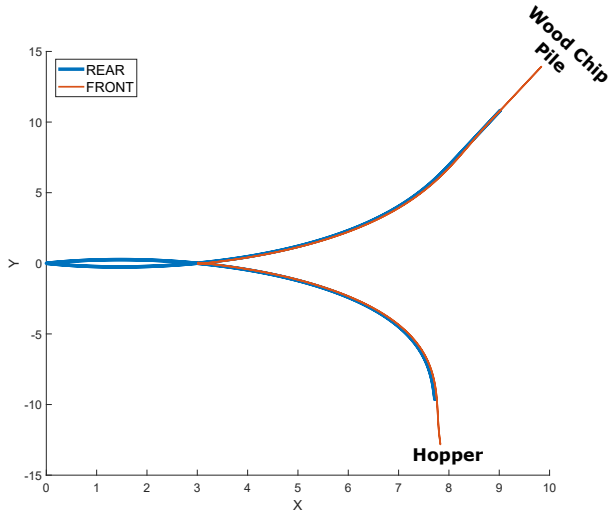


Fig. 4. Position

Figure 5 shows the heading and steering rates, which constitute the input of the articulated vehicle, as well as the lift

and tilt angles which are, in turn, the input of the loader arm.

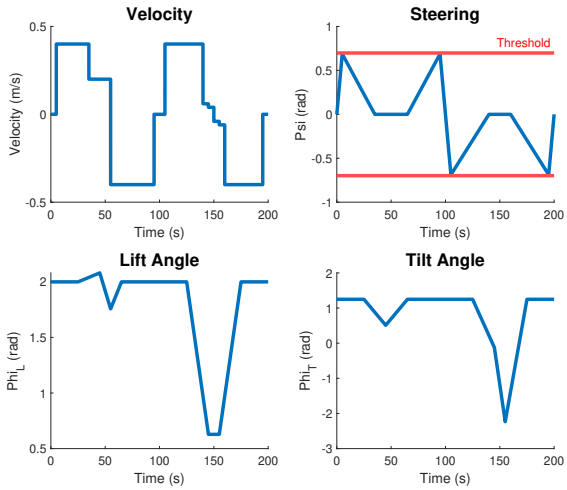


Fig. 5. Input

Finally, Fig. 7 shows the motion of the articulated vehicle (left) and the loader arm (right).

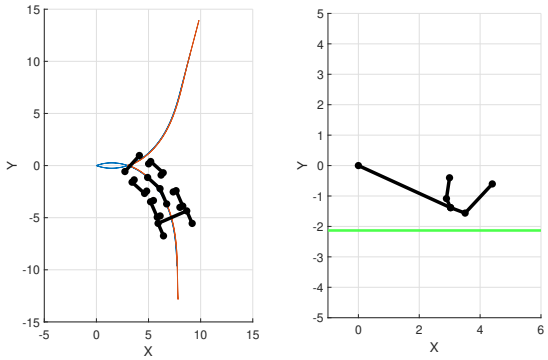


Fig. 6. Animation Frame

5. GAZEBO & ROS2 SIMULATION RESULTS

One can notice easily that although Matlab/Simulink is an excellent tool to build controllers, it might not be not the best when it comes 3D animation. Moreover, the dynamics was not covered yet, which might not be essential for a quick virtual proof-of-concept. However, since the goal is to end up with a digital twin of the wheel loader, an accurate simulation involving dynamics becomes a MUST. For that, we have modelled the wheel loader in Gazebo and ROS2 as shown in Fig. 7.

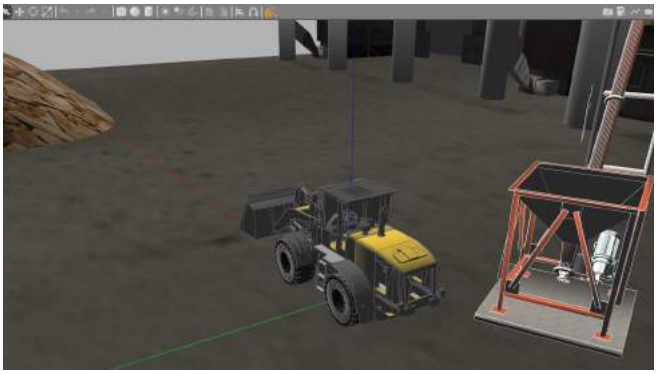


Fig. 7. Wheel loader in Gazebo

The Gazebo model contains seven joints, namely: center articulation steering, two rear wheels, two front wheels, arm and bucket. All joints are specified as revolute joints and thus will be controlled in angular velocities. A key consideration are the four wheel joints, as the articulated vehicle kinematics model elaborated in eq.(3) did not cover them all. The rear wheels, as mentioned above, can be associated with an equivalent center, leading to:

$$\dot{\theta}_{R1} = \dot{\theta}_{R2} = \dot{\theta} \tag{4}$$

While the front wheels must include compensation for the steering angle of the center articulation joint, therefore, the heading speed of the front chassis, v_f is:

$$v_F = v\cos\psi - a_R\dot{\varphi}_R\sin\psi \tag{5}$$

and the front wheels' rates are

$$\dot{\theta}_{F1} = \dot{\theta}_{F2} = v_F/r_w \tag{6}$$

The control algorithms developed in Section 4 will be maintained in Matlab/Simulink. The Gazebo model runs on Linux/Ubuntu machine. Parameters for dynamics have been calculated and incorporated into the Gazebo model. RT-LAB/Orchestra, a publisher-subscriber framework for co-simulation, is used to connect both entities across UDP. The reader can visualize the simulation here:
https://www.youtube.com/watch?v=UK8VCNPgSsI

6. CONCLUSIONS AND FUTURE WORK

As it is shown in the video, the bucket is kept empty because the contact dynamics, when loading and unloading wood chips, is not elaborated yet. This constitutes a major milestone in terms of future work. After that, the integration of virtual sensors is targeted to update in real-time the loader trajectory, which is pre-defined at the time being. Once we cover these two items, the kernel of the MIL will be completed and therefore we can move towards other steps, i.e., SIL, HIL and VIL.

7. REFERENCES

[1] Jonatan Björkman, "Control of an autonomous wheel loader," M.S. thesis, Lund University, Lund, Sweden, Sepember 2008.

[2] Etienne Abgrall, "Vision system for autonomous wheel loader," M.S. thesis, Royal Institute of Technology, Stockholm, Sweden, 2009.

[3] Robin Lilja, "A localisation and navigation system for an autonomous wheel loader," M.S. thesis, Mälardalen University, Västerås, Sweden, January 2011.

[4] Anders Bergdahl, "Autononmous bucket emptying on hauler," M.S. thesis, Linköpings Universitet, Linköping, Sweden, August 2011.

[5] Jonatan Blom, "Autonomous hauler loading," M.S. thesis, Mälaralen University, Västerås, Sweden, 2013.

[6] B.J. Dragt, F.R. Camisani-Calzolari, and I.K. Craig, "An overview of the automation of load-haul-dump vehicles in an undeground mining environment," *ELSEVIER IFAC Publications*, 2005.

[7] Shigeru Sarata, Noriho Koyachi, Takashi Tubouchi, Hisashi Osumi, Masamitu Kurisu, and Kazuhiro Sugawara, "Development of autonomous system for loading operation by wheel loader," *ISARC*, 2006.

[8] Jorge Angeles, *Fundamentals of Robotic Mechanical System: Theory, Methods, and Algorithms*, $4^{th} Edition, Springer, 2014$.

RIVER FLOW PATH CONTROL WITH REINFORCEMENT LEARNING

Dongqi LIU†, Yutaka NAITO†, Chen ZHANG†, Shogo MURAMATSU‡,
Hiroyasu YASUDA∗, Kiyoshi HAYASAKA¶, and Yu OTAKE§

†Graduate School of Sci. & Tech., Niigata Univ., Japan, ‡Faculty of Eng., Niigata Univ., Japan,
∗Research Inst. for Natural Hazard & Disaster Recovery, Niigata Univ., Japan,
¶Faculty of Sci., Niigata Univ., Japan, §School of Eng., Tohoku Univ., Japan

ABSTRACT

In this study, a cyber-physical system (CPS) for river flow path control is proposed using reinforcement learning. Recently, there has been a frequent occurrence of river flooding due to heavy rains, resulting in serious economic losses and victims. One cause of river flooding is the meandering due to the river bed growing and flow path change. As a mean of avoiding the meandering, river groynes can be used to regularize the flow. However, the mechanism of the flow path growing, and its optimal control is unclear. Therefore, in this study, a dynamic flow path control system is proposed using a data-driven approach to solve the problem at once. As a data-driven approach, reinforcement learning is adopted. The proposed system is designed to control meandering by adaptively deforming and moving the groynes with the reward of the flow path health. The effectiveness of the proposed flow path control system is verified through a simulation of the river model.

Index Terms— Reinforcement learning, Artificial intelligence, Cyber-physical system, Structural health monitoring

1. INTRODUCTION

In recent years, there has been a frequent occurence of river disasters caused by torrential rains, resulting in economic losses and victims. In Japan, the torrential rains in July 2020 and Typhoon No.10 in Sept. 2020 caused several river flooding and inundations, which requires urgent measures to mitigate the damage.

One of the causes of river flooding is the meandering due to changes in the flow path of the river channel [1]. The water-flow moves the river bed, creating shoals and meanders in the river. The shoals affect the flow path health even in ordinary conditions, contributing to the meandering river. Precipitation causes changes in the amount of flowing water, which tends to cause rapid flow in the transverse direction, resulting in overflow and erosion of the banks [2, 3]. To prevent and mitigate river flooding, it is necessary to elucidate the mechanism of the flow path change and control the path changes in the river channel. However, the actual flow path is diverse and difficult to theoretically describe. Thus, we have developed an indoor water-flow experiment system and try to understand and analyze the meandering mechanism with a data-driven approach [4, 5].

For regularizing the flow path, Umeki *et al.* developed a construction method of the artificial variable-width channel (AVWC), which can produce various flows by placing river groynes on both banks of the river [6]. In the formation of AVWC, the flow control can be achieved through an appropriate arrangement of static river groynes. However, an adaptive regularization of the flow path is demanded to control dynamically changing river states. Thus, it is necessary to deform or move the river groynes according to the river conditions.

In this study, we propose to develop a cyber-physical system (CPS) to adaptively control the river flow path, which keeps the healthy river state and helps us understand the meandering mechanism [7]. The following list summarizes the contribution of this study:

- proposal of a CPS concept to maintain and restore river channel health,
- development of a simulation model of a data-driven flow path control system, and
- validation of the significance of reinforcement learning for CPS [8].

This paper is organized as follows. Section 2 reviews the AVWC construction, while Section 3 shows our concept on the proposed CPS and a brief explanation of reinforcement learning. In Section 4, a simulation model is developed to validate the significance of our CPS concept. In Section 5, a performance evaluation is conducted to verify the effect of the proposed method. Finally, Section 6 is the conclusion.

2. ARTIFICIAL VARIABLE-WIDTH CHANNEL

A groyne is a rigid hydraulic structure that protects the shore of an ocean or the bank of a river from the flow of water and restricts the movement of sediment. River groynes can be used to control the velocity of water and to intercept or deflect the flow. Based on the analysis of experiment and simulation, groyne along the river bank helps reduce the river flow's speed and accelerate sedimentation [9, 10]. It is well known that alternating sandbars are generated and developed in rivers. However, the alternating sandbars cause problems in flood control and the environment. Natural variable-width channels often show the suppression of the generation and development of alternating sandbars. Taking a lesson from the function of nature, Umeki *et al.* have developed the AVWC method based on the river groyne construction techniques [6]. The AVWC is constructed by installing groynes on the banks on both sides of the river to create spatial heterogeneity in the flow and reduce the movement of the river bed. This method can artificially create a variety of flows and contribute to the maintenance of diverse ecosystems. The effectiveness of the AVWC method has been validated in real rivers.

In the case of small and medium-sized rivers, localized flood control and river environment measures are required. Since the river channels are often developed in a straight line, AVWC have the

This work was supported by JSPS KAKENHI Grant Numbers 20K20543, JP19K22026, JP19H04135.

Fig. 1. Example of AVWC in Hayade River, Agano River system, Japan (Provided by MILT of Japan)

advantage of being installed readily. Fig. 1 shows an example of AVWC in the Hayade River, Niigata, Japan. It has been confirmed that AVWC has good durability and does not change the shape of the river bank. Fig. 2 shows an example of AVWC using the indoor experimental setup [4]. In Fig. 2 (a), meandering is observed in the channel without the groynes on the river bank. On the other hand, as shown in Fig. 2 (b), in the channel with groynes on both banks, a straight flow is maintained compared to Fig. 2 (a). It is possible to create a healthy river environment by installing AVWC. However, static groyne placement is difficult to cope with actual rivers with constantly changing flow path. Therefore, a dynamic control system of AVWC is demanded.

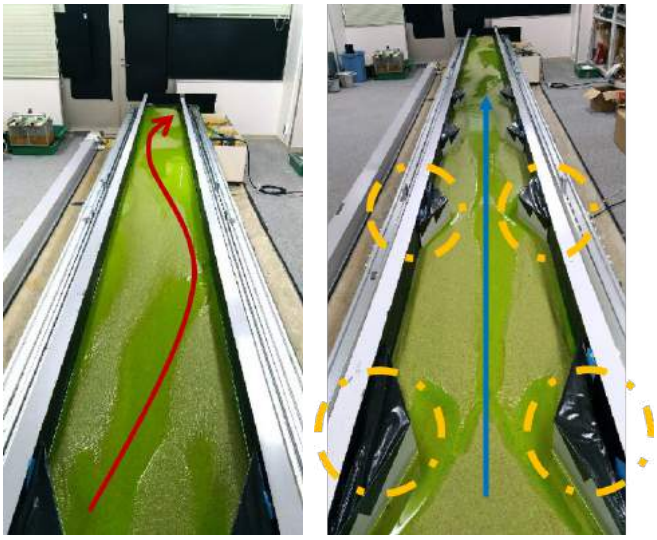
3. SYSTEM CONCEPT

Let us suggest a CPS concept using reinforcement learning as a data-driven method to maintain and restore river channel health on the unclear mechanism of flow fluctuation. This section provides an overview of the flow path control system and reinforcement learning.

3.1. Flow Path Control System

With the development of sensing, networking, and computing, we have been able to implement CPS in society, and its application is expected to realize some of the sustainable development goals (SDGs) [7]. We constructed a CPS by connecting the physical world to cyber space to autonomously solve certain social problems. In this study, we propose a CPS that controls a constantly fluctuating flow path of the river. Fig. 3 shows the configuration of our proposed CPS with a data-driven approach. As one of the data-driven methods, reinforcement learning has been applied in various areas such as automated adversary emulation for cyber-attacks [11, 12]. In this study, we propose to apply reinforcement learning to the river flow path control CPS and evaluate it.

Our proposed system consists of sensors such as a camera or radar that measures a river health index which is sent out to a server. The server uses the index as a reward to determine the optimal form, or placement of river groynes with reinforcement learning, or both, and sends control signals to the actuators for the deformation or translation, or both. The actuators are expected to autonomously deform and move groynes to constitute an optimal AVWC in the river to maintain a healthy environment.



(a) (b)

Fig. 2. Experimental results of AVWC using an indoor experimental setup. (a) without the river groynes, (b) with the river groynes (structures in black). The flow path specification: Length: 12 m, Width: 0.45 m, Gradient: 1/200, Flow rate: 2.0 ℓ/sec.

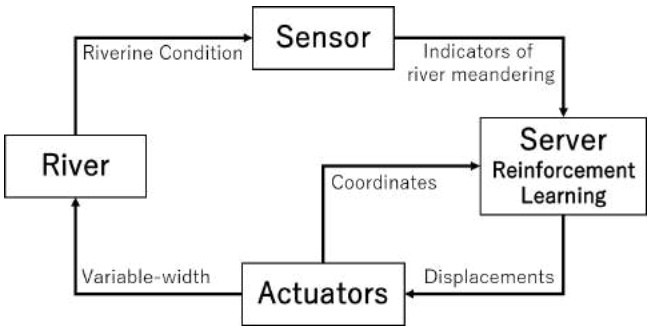


Fig. 3. Configuration of flow path control system

3.2. Reinforcement Learning

In recent years, there has been a significant improvement in determining, predicting, and working with unknown data in many fields where the mechanisms are unclear due to artificial intelligence (AI). These areas include robot control in manufacturing, automated driving technology, and opponents in games like AlphaGo. One of the fundamental technologies for AI is machine learning [8, 13], which gives us methods for solving problems such as prediction and classification through extracting statistical relationships on big data. The method can be broadly classified into three types: supervised learning, unsupervised learning, and reinforcement learning. Reinforcement learning is a method for learning "control" to "decide" the optimal "action" based on "recognition" of the situation, whereas the supervised and unsupervised learning approaches are based on "cognition."

Reinforcement learning is based on Bellman equation [14]. Let

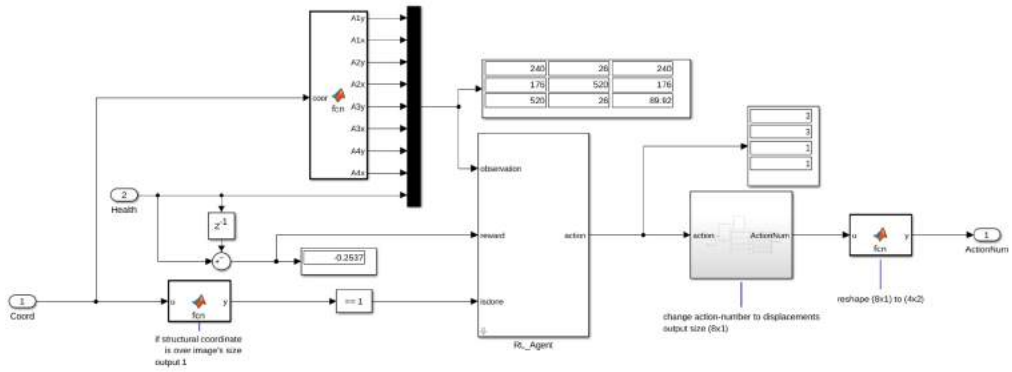


Fig. 4. Simulation model of the flow path control system with Deep Q-learning Network (DQN). The "Coord" terminal receives the coordinates of the river groynes and determines whether the action range has been exceeded. The "Health" terminal receives the health index h and combines it with the groynes coordinates to form an observation signal, input to the "observation" of the agent block. The change of the health index is input to the "reward." When the change exceeds the operating range, the logical decision "true" will be input to "isdone."

$\mathcal{S}$ and $\mathcal{A}$ be a state and action space, respectively. Under a stochastic policy $\pi(\cdot|s)\colon \mathcal{A} \to [0,1]$ given $s \in \mathcal{S}$, the expected value of the policy is represented by

$$v_\pi(s) = \sum_{a\in\mathcal{A}} \pi(a|s) \sum_{s'\in\mathcal{S}} p(s'|s,a)[r(s,a,s') + \gamma v_\pi(s')],\ s \in \mathcal{S} \tag{1}$$

where $r(s,a,s')$ is a reward from the state s to the immediate state s' on action a, $\gamma \in [0,1]$ is a discount factor that trades off the importance of immediate and later rewards, and $p(s'|s,a)$ is a state transition probability from the state s to the immediate state s' on action a. To determine optimal actions, the maximum value is calculated. As a model that uses Bellman equation, it is worthy to note that Markov Decision Process is used in the process. However, it is well known that building a probability tree over the whole problem space for large problems is difficult. In this study, we use the Deep Q-learning Network (DQN) algorithm to decide the optimal action a [15]. A Q-learning is a technique for choosing an action a over a given state s with no transition model to future states. In the Q-learning, a Q-value defined by

$$Q(s,a) \leftarrow Q(s,a) + \alpha\delta(s',a'), \tag{2a}$$

$$\delta(s',a') = r(s) + \gamma \max Q(s',a') - Q(s,a) \tag{2b}$$

is used, where α is a scalar step size. We used DQN that employs deep learning for estimating the Q-value. Since repeated updates in the same network is prone to overestimate the value of action a, a double DQN technique has been developed, where a network for selecting the action a and another network for calculating the target value $\gamma \max Q(s',a')$ of the function $Q(s,a)$ are connected. In addition, the negative effects of parameter updates are reduced and the performance improvement is expected [16].

4. SIMULATION MODEL OF PROPOSED CPS

Model in the Loop Simulation (MILS), i.e., a virtual phase of model-based development of a control system, is beneficial for debugging in the early design stage [17]. If a simulation can reproduce actual behavior of the control target in which a control model can be developed by MILS. The control system can then be applied to a real target. This development approach accelerates the construction of the entire real system. We propose to adopt reinforcement learning to adaptively control the flow path of a river, develop a fully virtual simulation model of the entire system to detect issues in the early development stage. Also, we replace the target model for control with an experimental setup of the indoor river model, i.e., to expand the model to Component in the Loop Simulation (CILS).

We modeled the configuration of the flow path control system shown in Fig. 3 with MATLAB/Simulink. Fig. 4 shows the MATLAB/Simulink model of the fully virtual simulation model of the entire system. For the reinforcement learning model, we developed a DQN network mentioned in Section 3 using the agent block of MATLAB Reinforcement Learning Toolbox. The agent block has three terminals "observation", "reward" and "isdone" as inputs, and one terminal "action" as an output. We assume four groynes in our simulation model. When the health index h is observed, eight scalar quantities of the current coordinates of positions in (y,x) of the four groynes are combined with the health index h and sent to the control system, where a signal of dimension 1×9 is input as the state observation. The change of the health index h is computed and input as the reward. We use the reward to update the Q-value of the determined action for the observed signal. At this point, the current action episode is stopped, and a new action learning episode will start when the structure coordinates exceed the action range. We now evaluate the performance of the reinforcement learning model in Section 5.

5. PERFORMANCE EVALUATION

In this study, we used MATLAB/Simulink R2020b for creating the simulation model of the entire system. Reinforcement Learning Toolbox and Deep Learning Toolbox were used to create a DQN model on MATLAB/Simulink. Table 1 summarizes the specification.

To evaluate the effectiveness of the proposed simulation model with reinforcement learning for the flow path control system, three river images of different meander cycles were used. Fig. 5 shows the simulation results, where blue area indicates river channel, the brown area indicates river soil and black area indicate river groynes. We simply defined the health index h by the ratio of the river soil area in the entire image. In this simulation, the most rewarding activity is to move the groynes into the bends of the river, since the river groynes are considered as river soil areas in the health index h calculation. It is worthy to note that the dynamics of the river soil

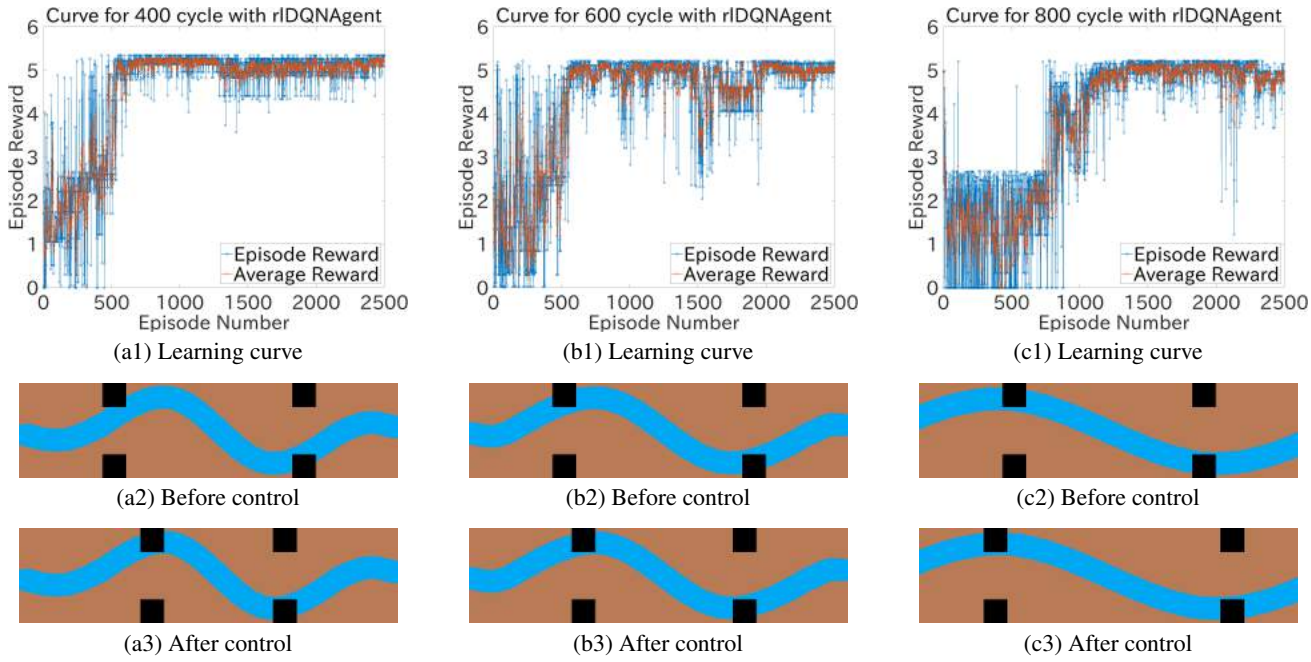


Fig. 5. Simulation results for river channel image of size 800×200 pixels. (a#) Results of meander cycle 400 pixels; (a1) Learning curve, (a2) Initial state of river groynes, and (a3) Optimal state of river groynes, where the health index h changes from 0.848 to 0.899. (b#) Results of meander cycle 600 pixels; (b1) Learning curve, (b2) Initial state of river groynes, and (b3) Optimal state of river groynes, where the health index h changes from 0.825 to 0.875. (c#) Results of meander cycle 800 pixels; (c1) Learning curve, (c2) Initial state of river groynes, and (c3) Optimal state of river groynes, where the health index h changes from 0.755 to 0.807

movement will be modeled in the future and the health index h will be refined by considering the degree of curvature. Also, note that the current control model can be used when replacing the target component and the health index in the entire CPS model. Table 2 summarizes the parameters of the DQN model. The simulation model learns 2500 episodes on random initial states of river groynes to achieve the maximum value within maximum steps of 20 per episode moving 20 pixels each step.

After learning, defining the initial state of river groynes to 200 and 600 pixels, the followings summarize the simulation results:

(a) Meander cycle of 400 pixels: The health index h changed from 0.848 to 0.899.

(b) Meander cycle of 600 pixels: The health index h changed from 0.825 to 0.875.

(c) Meander cycle of 800 pixels: The health index h changed from 0.755 to 0.807.

In all cases, river groynes were moved to the bent parts of the river images as expected.

From Figs. 5 (a1), (b1), and (c1), we observed that every learning curve converges to almost the highest value, even though there were some errors on the groynes' movement. In summary, the simulation results for the three different meander cycles show that the flow path control system can effectively works with reinforcement learning so that the river groynes can autonomously move to the expected positions and the health indices increased.

Table 1. Experimental specifications

OS	Ubuntu 18.04 LTS
Language	MATLAB/Simulink R2020b
Toolboxes	Reinforcement Learning Toolbox Deep Learning Toolbox

Table 2. Parameters of the DQN model

Discount Factor γ	0.9
Learning Rate	0.001
Maximum Number of Episode	2500
Maximum Steps per Episode	20

6. CONCLUSION

In this study, we proposed a CPS concept for controlling river channels using reinforcement learning as the controller. We developed a simulation model and verified the significance of the proposed system by evaluating river images of different meander cycles. It was confirmed that the proposed configuration can be used to maintain the river channel, and the reinforcement learning is effective for the control. In the future, we will replace the static picture in simulation with a dynamic motion model of water and sediment, define a more realistic health index, and conduct CILS with river experimental setup as the control target. It is also hoped that machine learning will help clarify the mechanism of meandering and the optimal water control method. Applications of reinforcement learning in disaster prevention could also be expected.

7. REFERENCES

[1] S. Yamaguchi, T. Kyuka, Y. Shimizu, N. Izumi, Y. Watanabe, and T. Iwasaki, "Influence of Sediment Dynamics in River Channel on the Channel Change (in Japanese)," *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*, vol. 74, no. 4, pp. I_1153–I_1158, 2018.

[2] T. Sumner, T. Inoue, and Y. Shimizu, "A Study of Sandbar Formation on Bedrock and Bedrock Erosional Morphology (in Japanese)," *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*, vol. 72, no. 4, pp. I_817–I_822, 2016.

[3] Y. Shimizu, K. Osada, and T. Takanashi, "Numerical Study on the Formation of Low-water Course in a Straight Channel with Alternate Bars (in Japanese)," *Proceedings of Hydraulic Engineering*, vol. 48, pp. 1027–1032, 2 2004.

[4] T. Hoshino, H. Yasuda, and M. Kurashiki, "Direct Measurement Method of Formation Process of Alternate Bars (in Japanese) ," *Journal of Japan Society of Civil Engineers, Ser. A2 (Applied Mechanics (AM))*, vol. 74, no. 1, pp. 63–74, 2018.

[5] Y. Kaneko, S. Muramatsu, H. Yasuda, K. Hayasaka, Y. Otake, S. Ono, and M. Yukawa, "Convolutional-sparse-coded Dynamic Mode Decomposition and Its Application to River State Estimation," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. 5 2019, vol. 2019-May, pp. 1872–1876, Institute of Electrical and Electronics Engineers Inc.

[6] K. Umeki, H. Yasuda, I. Ono, Y. Hosaka, K. Shimizu, and K. Kuroishi, "Validiation of Funtion Both of Flood Control and Envaironment Protection on Artificial Variable Width Channel in the Hayade River (in Japanese) (to be published)," *Advances in River Engineering*, vol. 26, 2021.

[7] J. Jamaludin and J. M. Rohani, "Cyber-Physical System (CPS): State of the art," in *2018 International Conference on Computing, Electronic and Electrical Engineering, ICE Cube 2018*. 1 2019, Institute of Electrical and Electronics Engineers Inc.

[8] K. Arulkumaran, M. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," 11 2017.

[9] S. Krishna Prasad, K.P. Indulekha, and K. Balan, "Analysis of Groyne Placement on Minimising River Bank Erosion," *Procedia Technology*, vol. 24, pp. 47–53, 1 2016.

[10] W. Prasetyo, P. T. Juwono, and D. Sisinggih, "Analysis on The Effect of Groyne Type Impermeable Placement on Sediment Distribution in Lariang River Bend," *Civil and Environmental Science Journal*, vol. 4, no. 1, pp. 43–061, 1 2021.

[11] X. Liu, H. Xu, W. Liao, and W. Yu, "Reinforcement learning for cyber-physical systems," in *Proceedings - IEEE International Conference on Industrial Internet Cloud, ICII 2019*. 11 2019, pp. 318–327, Institute of Electrical and Electronics Engineers Inc.

[12] A. Bhattacharya, T. Ramachandran, S. Banik, C. P. Dowling, and S. D. Bopardikar, "Automated Adversary Emulation for Cyber-Physical Systems via Reinforcement Learning," in *Proceedings - 2020 IEEE International Conference on Intelligence and Security Informatics, ISI 2020*. 11 2020, Institute of Electrical and Electronics Engineers Inc.

[13] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2 2015.

[14] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," 2012.

[15] N. Justesen, P. Bontrager, J. Togelius, and S. Risi, "Deep learning for video game playing," *IEEE Transactions on Games*, vol. 12, no. 1, pp. 1–20, 3 2020.

[16] H. van Hasselt, A. Guez, and D. Silver, "Deep Reinforcement Learning with Double Q-learning," *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, pp. 2094–2100, 9 2015.

[17] J. Kapinski, J. V. Deshmukh, X. Jin, H. Ito, and K. Butts, "Simulation-Based Approaches for Verification of Embedded Control Systems: An Overview of Traditional and Advanced Modeling, Testing, and Verification Techniques," *IEEE Control Systems*, vol. 36, no. 6, pp. 45–64, 12 2016.

IMPROVING AUTOMATED SEARCH FOR UNDERWATER THREATS USING MULTISTATIC SENSOR FIELDS BY INCORPORATING UNCONFIRMED TRACK INFORMATION

D. Angley, S. Mehrkanoon, B. Moran

School of Engineering,
The University of Melbourne,
Parkville, VIC 3010

C. Gilliam

School of Engineering,
RMIT University,
Melbourne, VIC 3000

S. Simakov

Maritime Division,
DSTG, Edinburgh,
SA 5111, Australia

ABSTRACT

Sonobuoy fields, comprising a network of sonar transmitters and receivers, are used to search for and track underwater targets. Although normally such fields are operated from a maritime patrol aircraft, automated scheduling and processing creates opportunities for employing them as autonomous sensor systems. The automated search mechanism considered in this work is controlled by modelling the presence of undetected threats in an Operational Area (OA) using a spatial probability density function (PDF), known as a threat map. The algorithm decides how to schedule waveform transmissions, known as pings, to efficiently search and clear the OA. A conventional approach is to update the threat map based on just the characteristics of the sonobuoy field and switch to a separate metric to track a target after track confirmation. In this study we address the phase when there are potential contacts which cannot yet be promoted to confirmed tracks. We develop a mechanism for probing the associated areas of interest while still remaining in the threat map driven search scheduling. To this end, we propose reinitialising the threat map after each transmission using an augmented PDF, where unconfirmed tracks are represented by weighted Gaussians. Simulations show that this approach significantly improves search performance, reducing the number of pings required to confirm a track, distance from a confirmed track to the target and the proportion of falsely confirmed tracks.

***Index Terms*—** Multi-static sonar, Sensor scheduling, Autonomous search.

1. INTRODUCTION

A sonobuoy is a compact deployable sonar system containing sonar transmitters and/or receivers that can be used to search for and track underwater targets. Multiple sonobuoys can be laid out in a sonobuoy field and be used cooperatively, with an operator or autonomous sensor management system deciding which sonobuoys should ping at any given time. The goal of the sonobuoy field is to search an Operational Area (OA) for an undetected threat; success is either the clearance of the area or the detection and accurate tracking of a target. In this paper, we focus on the search mode of operation. Once the presence of a target has been confirmed then the field switches to a tracking mode, where the focus is on high accuracy location and tracking of the target. A challenging aspect, however, is the confirmation of target; each transmission by the sonobuoy field results in many detections that may or may not relate to a target. To avoid false tracks, a track is confirmed once there is a high probability that the track is associated with a target. This confirmation process requires multiple detections to be associated across multiple pings, and hence coordination in the search mode to ensure detection opportunities. Accordingly, our focus is on optimising the search mode to reduce the number of pings (i.e. time) required to confirm a track and clear an area.

A threat map is used to model the spatial PDF of an existing, but undetected target [1–7]. This spatial PDF is updated as sonobuoys in the field transmit waveforms and process returns, lowering the probability that an undetected target could exist in areas near the transmitting sonobuoy according to the probability of detection. The PDF is also updated at each time step to take into account possible kinematics of the target. The threat map can be used as a planning tool, helping the operator to decide both where to place the sonobuoys and when each sonobuoy should transmit a waveform in order to effectively search an operational area. Section 2 details multiple approaches to calculating the threat map found in literature.

The existing methods of calculating a threat map are generally focused on modelling detections and detection probability, but in practice the decision to switch to tracking mode is based on more than just a single detection. The sonar environment is usually noisy and spurious detections are common. The system cannot switch to tracking mode for every detection, because this would waste time trying to track a target that does not exist while a true target continues to operate in another area. Detections are therefore processed by a tracker, which requires multiple consistent detections to confirm the presence of a target. Detections initially create unconfirmed tracks, with a low probability of being associated with a true target, and are either promoted to confirmed tracks or discarded based on subsequent detections.

In our previous work using threat maps to optimise sonobuoy search [8], we observed a mismatch between detecting and confirming a track, which could lead to suboptimal performance. A sonobuoy might ping and clear an area of the threat map, but if this was not followed up with subsequent pings in the same area then the tracker might discard the unconfirmed tracks by the time the area was revisited. In this paper we explore the impact of including the unconfirmed tracks from the tracker into the threat map in order to make more effective search decisions, such as repeatedly pinging in the same area in order to confirm or discard unconfirmed tracks. This is a novel approach that also has the advantage of being easy to retrofit to existing systems, taking advantage of their existing tracking algorithms. It moves application of the threat map from the planning of the mission to its execution, where additional information from detections can be used to schedule sonobuoy transmissions.

2. THREAT MAP

The threat map is a representation of the spatial PDF of an existing but undetected target. There are two main approaches in the literature for modelling undetected targets.

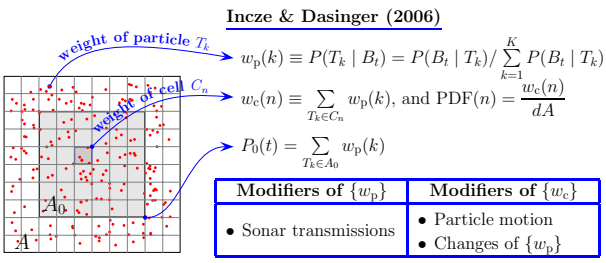


Fig. 1. The Incze & Dasinger [5, 6] approach to modelling the threat map. A large number of virtual targets (red dots) are distributed over area A. The particle weight (w_p) is calculated as the probability that the target is realised as virtual target k (event T_k) given that it hasn't been detected yet (event B_t), divided by the sum of probabilities for all K particles. The weight of a cell (w_c) is the sum the weights of all the particles within it, and the PDF of the target location is the cell weight divided by the cell area (dA). An area of interest (A_0) is shown, and the probability that the target is in that area (P_0) can be calculated by summing the weights of the particles it contains.

Krout et al. [1–4] model the probability of target presence using a regular grid of cells, and allow for target movement using a drift and diffusion model on this grid. The calculated parameter $P_{(i,j),k}$ is the probability of target presence in a cell (i, j) at step k, which is incremented at each time step or whenever the sonar transmission occurs. The amount by which a sonar transmission changes $P_{(i,j),k}$ depends on the probability of detection in the cell, while value $P_{(i,j),k+1}$ after the time step update is obtained using $P_{(s,q),k}$ from adjacent cells (s, q) by application of a 3×3 spatial Fokker-Planck (FP) filter which depends on the "drift" and "diffusion" coefficients associated with the considered target motion scenario. We used this technique in our previous work, which considered the tradeoff between tracking and search when scheduling transmissions in sonobuoy fields [8]. One of the limitations of the approach based on the FP filtering is that, for more complex scenarios, the translation of the target motion assumptions into drift and diffusion coefficients is not always straightforward. Using arbitrary values for these coefficients may produce an invalid filter. Some scenarios may require a higher-order finite-difference approximation and a larger size (5×5) of the FP filter. Note also that the fact that the probability of presence does not integrate to 1 implies that its transformation into a probability density function requires some form of rescaling. Such a rescaling procedure and an interpretation of its result are yet to be discussed in the published literature.

Incze and Dasinger [5, 6] call their model of the threat map a "Threat Density Probability Map". They use a Monte Carlo approach in which the target can be realised as one of the virtual targets from a large set of virtual targets initialised and propagated in accordance with the scenario of interest. The probability of target presence in an area given that it remains undetected as the sonobuoys transmit is calculated using Bayes's Theorem and a model of detection probability, as well as the target motion assumptions. This probability is normalised to sum to 1 over all target locations considered, modelling the assumption that a single target exists somewhere. Figure 1 details this approach.

The Incze and Dasinger method is computationally expensive because of the large number of virtual targets. However, Simakov and Fletcher [7] implement this model on the GPU, allowing for tens of thousands of virtual targets to be used. We use this approach to calculate the threat map in this paper, and Figure 2 shows an example of how this threat map evolves as sonobuoys ping.

A limitation of these approaches is that they do not take into account detection information available during operation. We now propose a method of incorporating this information into the threat map in order to reduce the number of pings required to confirm a target.

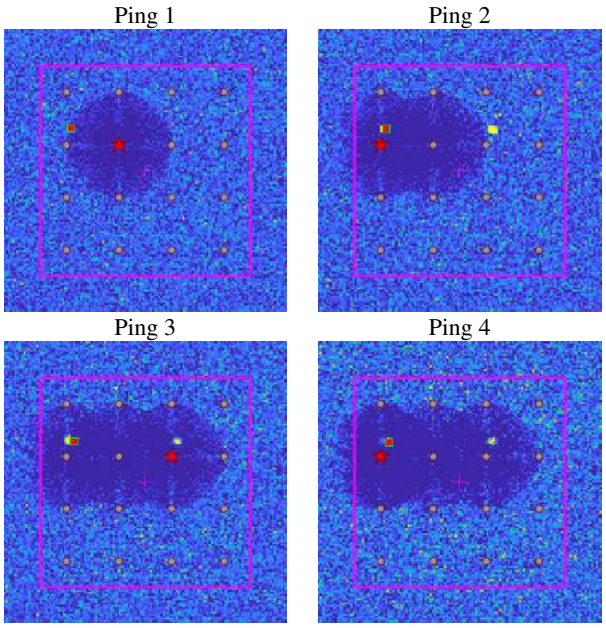


Fig. 2. Example showing how the threat map evolves as sonobuoys ping over four consecutive pings. The colour represents the probability density of target presence, with darker areas indicating lower probability. The sixteen grey circles show the sonobuoys, which are all both transmitters and receivers. The pink square shows the operational area which is being searched. The red square with a green outline shows the target location. The red star on a sonobuoy indicates that it has just emitted a waveform, which clears a roughly circular area around it by lowering the probability of target presence. The bright yellow blobs show unconfirmed tracks which have been incorporated into the threat map.

3. INCORPORATING UNCONFIRMED TRACKS INTO THE THREAT MAP

To overcome the limitations discussed previously, we now present our approach to incorporating unconfirmed target information into the threat map. For each sonar transmission, a small number of detections are received, which are used to update existing confirmed tracks or initiate a new unconfirmed track. The location $\mathbf{x}_k \equiv [x_k, y_k]^T$ and covariance matrix $\mathbf{C}_{xy}^k$ of each unconfirmed track ($k = 1, \ldots, K_t$), which are outputs from the tracker, were used to form the associated 2D Gaussian corrections to the evolving Threat Map PDF $p(\mathbf{x}, t)$

$$p(\mathbf{x}, t) \mapsto \gamma_0 p(\mathbf{x}, t) + \sum_{k=1}^{K_t} \gamma_k G(\mathbf{x} - \mathbf{x}_k; \mathbf{C}_{xy}^k) \qquad (1)$$

where

$$G(\mathbf{x}; \mathbf{C}_{xy}) \equiv \frac{\exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{C}_{xy}^{-1} \mathbf{x}\right)}{2\pi\sqrt{\det \mathbf{C}_{xy}}}$$

and $\sum_{k=0}^{K_t} \gamma_k = 1$. After each sonar transmission we first apply a standard threat map update [7], which accounts for ping-induced reduction of weights of virtual targets. In the simulations discussed in this paper, γ_k used in (1) had the following values: γ_0 fixed and $\gamma_k = (1 - \gamma_0)/K_t$ ($k > 0$).

Next we incorporate unconfirmed tracks into the resulting $p(\mathbf{x}, t)$ by adding weighted Gaussian corrections. This also uses

a reinitialisation procedure which produces a set of K_p equally-weighted particles $\{\mathbf{x}_\mathrm{p}(k),\ w_\mathrm{p}(k) = 1/K_\mathrm{p}\}$ spread over the cells of the threat map. Figure 3 illustrates calculation of the cell index for particle k. Here $U(0,1)$ is the uniform distribution in the interval $(0,1)$. We use the inversion method [9, Ch. 3]. The probability that a particle is placed into cell n is $w_\mathrm{c}(n)$, where the cell weights $\{w_\mathrm{c}(m)\}$ are obtained by integrating the post-ping PDF $p(\mathbf{x},t)$ in the respective cells. Once the cell for particle k has been identified, $\mathbf{x}_\mathrm{p}(k)$ is obtained by drawing from a uniform distribution defined in the cell. This process effectively imbues the threat map with a memory of all the previous detection information.

The repeated reinitialisation can result in clumping of particles into some cells at the expense of depopulation of the adjacent cells. Replacing pseudo-random generation used for production of $r_k \in U(0,1)$ during calculation of the particle cell index with a quasi-random approach (e.g. Sobol, or scrambled Sobol generators [10]) resolves this issue.

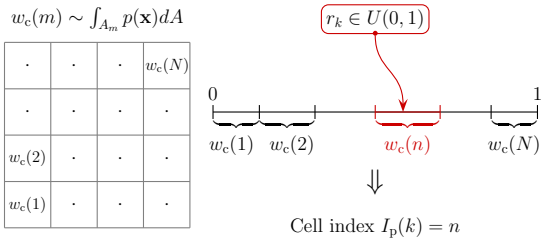


Fig. 3. Reinitialisation procedure: calculation of cell-index of particle k.

4. SCHEDULING MULTISTATIC SONOBUOY FIELDS

The key elements required to schedule a sonobuoy field are introduced here.

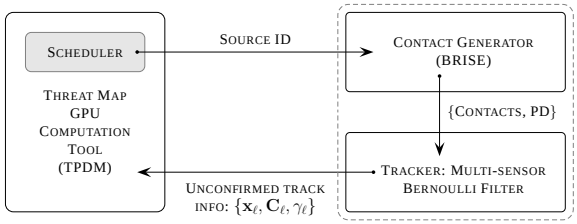
4.1. Sensor Scheduling Algorithm

Sensor scheduling in this sonobuoy scenario means choosing which sonobuoy to emit a waveform from in each ping interval T_p. In this paper we consider three scheduling algorithms: random, raster and lookahead scheduling. The random scheduler picks a uniformly random sonobuoy to transmit in each interval. The raster scheduler uses a predefined raster pattern of sonobuoys to ping from that loops until the target is found. The lookahead scheduler is a myopic scheduler that picks the sonobuoy that minimises the probability of target presence in the area of interest, A_0, after the next ping, as modelled by the threat map. Note, as discussed in the introduction, we only consider scheduling to optimise the search for an unknown target; the tracking and localisation of a confirmed target requires different optimisation criteria [8, 11].

We apply the lookahead scheduler to the threat map evolved in two different modes: conventional, as described by Figure 1, and our method of incorporating unconfirmed track information, as described in Section 3. Figure 4 shows a diagram of the simulation setup and summarises how the lookahead scheduler works when the latter mode is employed.

4.2. Measurement Simulation

We use the BRISE (Bistatic Range Independent Signal Excess) [12] simulation environment to simulate sonobuoy measurements. BRISE



Adaptive (lookahead) ping selection workflow:

1. Before a transmission, probe different sources m to examine the difference between $p(\mathbf{x},t)$ and the post-ping $p_m(\mathbf{x})$
2. Select source ID s resulting in maximal reduction of the Probability of Presence P_0
3. Set post-ping PDF: $p(\mathbf{x},t) = p_s(\mathbf{x})$, where $s = \operatorname{indmax}_m \left(P_0[p(x,t)] - P_0[p_m(\mathbf{x})]\right)$
4. Contact generator uses the actual target(s) and the source ID s to generate contacts
5. The tracker uses the contacts and the source ID s to produce $\{\mathbf{x}_\ell, \mathbf{C}_\ell, \gamma_\ell\}$
6. Update post-ping PDF: $p(\mathbf{x},t) \to \gamma_0 p(\mathbf{x},t) + \sum_{\ell=1}^{L} \gamma_\ell G_\ell(\mathbf{x} - \mathbf{x}_\ell; \mathbf{C}_\ell)$
7. Propagate virtual targets and evolve $p(\mathbf{x},t)$ up to the time of the next transmission

Fig. 4. Simulator setup, showing interaction between the threat map GPU computation tool, the contact generator, and the tracker in the case when the lookahead scheduler is applied to a threat map augmented by unconfirmed tracks.

uses lookup tables containing signal excess data, precomputed using the Gaussian ray bundle eigenray propagation model [13], to calculate a signal-to-noise ratio (SNR) for each target and produce measurements of the bistatic range and bearing, along with false alarms [14]. The simulations in this paper all use a linear frequency modulation (LFM) waveform centred at 2 kHz with a transmission duration of 2 seconds and bandwidth of 200 Hz.

4.3. Tracker

In order to provide unconfirmed tracks to combine with the threat map, we use a multi-target tracking algorithm [14, 15] that has been extensively evaluated for tracking targets in similar sonobuoy scenarios [8, 11, 16, 17] and found to perform robustly. The tracker uses the multi-sensor Bernoulli filter [18, 19], the optimal Bayesian multi-sensor filter for a single target, and applies the linear-multitarget paradigm [20] to extend it to track multiple targets. In particular, we use the Gaussian mixture model implementation of this algorithm. The tracker estimates a probability of target existence for each track, along with a track status that indicates if the track is confirmed (i.e. high probability of being associated with a true target) or unconfirmed.

5. RESULTS

In order to evaluate the performance of the system shown in Figure 4 and demonstrate the advantage of using unconfirmed track information in the threat map, we used simulations of the scenario shown in Figure 5. This scenario has 16 sonobuoys, arranged in a 4×4 grid and spaced 20 km apart. In each simulation, the target starts at a random position uniformly distributed in the 5 km neighbourhood of the boundary of the operational area and moves towards the centre of the field. The scenario is run until the tracker reports its first confirmed track, at which point the field would switch modes from search to tracking in a practical scenario. The performance metrics gathered are the number of pings until the track is confirmed, the proportion of falsely confirmed tracks, and the distance between the target location and the confirmed track. A falsely confirmed track is

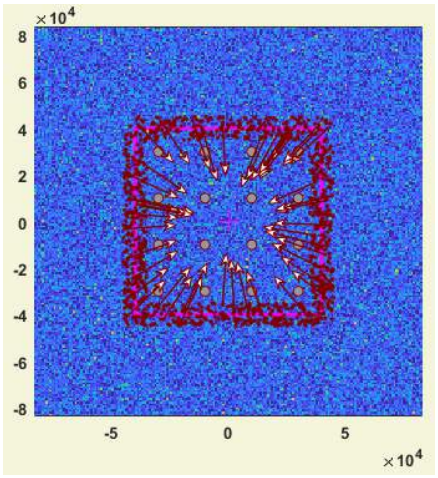


Fig. 5. Sonobuoy scenario, with 16 sonobuoys (grey circles) laid out in a grid spaced 20 km apart. The red dots show sample target trajectories, starting in the neighbourhood of the boundary of the operational area and moving towards the centre. The pink square shows A_0, the area of interest, which is a square area from $(-40, -40)$ km to $(40, 40)$ km.

defined as a confirmed track position estimate that is more than 1 km from the true target position, and these are excluded from the mean distance calculation.

Simulations were performed for the four different scheduling methods detailed in Section 4.1, which are labelled Raster (predefined sweeping pattern), Random (uniform random selection of sonobuoy), Lookahead (lookahead scheduling without incorporating unconfirmed tracks) and Unconfirmed (lookahead scheduling incorporating unconfirmed tracks). 1,000 Monte-Carlo simulations were performed for all methods and for various values of γ_0.

Figure 6 shows the results from these simulations. The three methods that do not use unconfirmed track information all show similar performance on the number of pings to confirm a track, taking a mean of 24.9 to 26 pings to confirm a track. Incorporating unconfirmed track information into the threat map reduced this number of pings significantly, down to 13.5 pings for $\gamma_0 = 0.93$. Importantly, the corresponding proportion of falsely confirmed tracks is also lower for the Unconfirmed method. Thus, the improvement in confirmation is due to the tracker confirming true tracks faster not false tracks. The mean distance between the target and the confirmed track at confirmation time was also lower for the Unconfirmed method, indicating that this method also improves track error at confirmation time. The performance of the Unconfirmed method was fairly stable across a wide range of γ_0 and converges with the performance of the Lookahead method as γ_0 approaches 1 as expected, because no unconfirmed track information is incorporated when $\gamma_0 = 1$.

Figure 7 shows 2D histograms of the proportion of pings emitted from each sonobuoy, with a layout matching Figure 2. Raster shows the characteristic vertical scanning pattern, from the bottom left to the top right. Random is uniformly distributed, as expected. Lookahead shows unbalanced usage of the field, focusing on the edges and corners once the centre of the field has been cleared because that is where a target will likely enter from. By contrast, Unconfirmed shows more balanced sonobuoy usage which is somewhere between Lookahead and Random, with a focus on the edges of the field but the unconfirmed tracks leading it to make greater usage of the other sonobuoys.

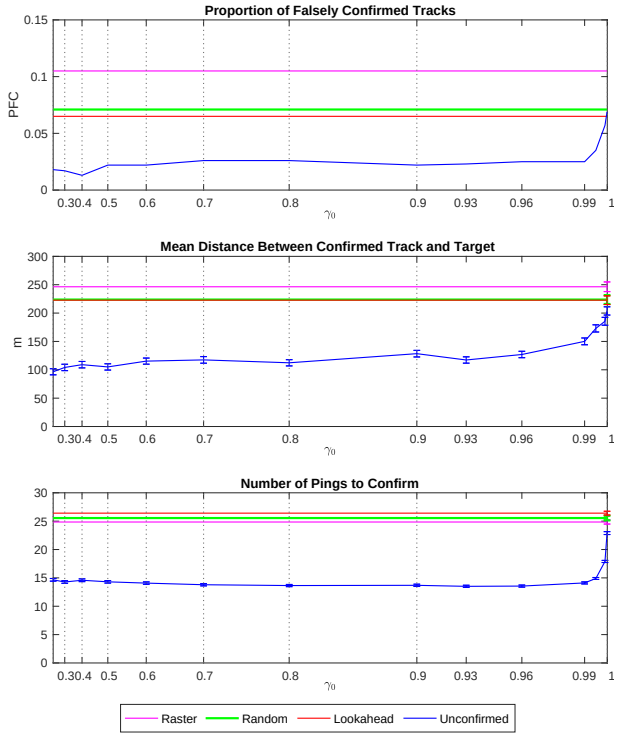


Fig. 6. Simulation results, showing the proportion of falsely confirmed tracks, mean distance between the confirmed track and the true target and the number of pings to confirm a target for the 4 methods. Performance is shown for a range of γ_0 values, with Unconfirmed being the only method that depends on this parameter. The error bars show plus or minus one standard deviation.

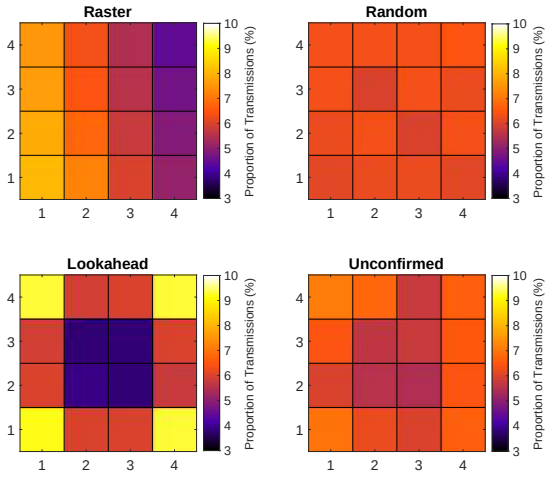


Fig. 7. 2D histogram of the proportion of pings transmitted by each sonobuoy of the 4×4 field in simulations of the four different scheduling methods.

6. CONCLUSION

In this paper we have presented a novel approach to improving autonomous search of an area of interest using a sonobuoy field. By incorporating unconfirmed track information from the tracker into the threat map, we demonstrate a significant increase in performance, reducing the number of transmissions required to confirm the presence of an underwater threat, such as a submarine or a sizeable UUV.

The confirmed tracks were also more accurate and there was a lower proportion of falsely confirmed tracks. This suggests that utilising the proposed technique in the search mode of multistatic systems could be valuable in practice, and our approach can take advantage of the existing tracking algorithms of such systems. Future work includes exploring alternative ways of combining unconfirmed tracks with the threat map.

7. REFERENCES

[1] D. W. Krout, M. A. El-Sharkawi, W. L. Fox, and M. U. Hazen, "Intelligent ping sequencing for multistatic sonar systems," in *Proceedings of the 9th International Conference on Information Fusion*. IEEE, 2006, pp. 1–6.

[2] D. W. Krout, W. L. Fox, and M. A. El-Sharkawi, "Probability of target presence for multistatic sonar ping sequencing," *Oceanic Engineering, IEEE Journal of*, vol. 34, no. 4, pp. 603–609, 2009.

[3] D. W. Krout, G. M. Anderson, E. Hanusa, and B. D. Jones, "Threat modeling for sensor optimization," in *2013 OCEANS-San Diego*. IEEE, 2013, pp. 1–4.

[4] D. W. Krout and T. Powers, "Sensor management for multistatics," in *17th International Conference on Information Fusion (FUSION)*. IEEE, 2014, pp. 1–6.

[5] B. I. Incze and S. B. Dasinger, "Revisiting measures of effectiveness in support of low-frequency, multistatic sonar search in the littoral battlespaces," Tech. Rep., Naval Undersea Warfare Center, Newport RI, 2000.

[6] B. I. Incze and S. B. Dasinger, "A bayesian method for managing uncertainties relating to distributed multistatic sensor search," in *Proceedings of the 9th International Conference on Information Fusion*. IEEE, 2006, pp. 1–7.

[7] S. Simakov and F. Fletcher, "GPU acceleration of threat map computation and application to selection of sonar field controls," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 1827–1831.

[8] C. Gilliam, B. Ristic, D. Angley, S. Suvorova, B. Moran, F. Fletcher, H. Gaetjens, and S. Simakov, "Scheduling of multistatic sonobuoy fields using multi-objective optimization," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 3206–3210.

[9] L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag, 1986.

[10] "cuRAND Library: Programming Guide," NVIDIA, PG–05328–050, July 2019, https://docs.nvidia.com/cuda/.

[11] C. Gilliam, D. Angley, S. Williams, B. Ristic, B. Moran, F. Fletcher, and S. Simakov, "Covariance cost functions for scheduling multistatic sonobuoy fields," in *2018 21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 1–8.

[12] S. Simakov, "Signal excess data and tools for multistatic sonar emulation," Tech. Report DSTO-TR-3026, DSTO, 2014.

[13] H. Weinberg and R. E. Keenan, "Gaussian ray bundles for modeling high-frequency propagation loss under shallow-water conditions," *The Journal of the Acoustical Society of America*, vol. 100, no. 3, pp. 1421–1431, 1996.

[14] B. Ristic, D. Angley, F. Fletcher, S. Simakov, H. Gaetjens, S. Suvorova, and B. Moran, "Bayesian multitarget tracker for multistatic sonobuoy systems," in *Proceedings of the 19th International Conference on Information Fusion*. IEEE, 2016, pp. 2171–2178.

[15] B. Ristic, D. Angley, S. Suvorova, B. Moran, F. Fletcher, H. Gaetjens, and S. Simakov, "Gaussian mixture multitarget–multisensor Bernoulli tracker for multistatic sonobuoy fields," *IET Radar, Sonar & Navigation*, vol. 11, no. 12, pp. 1790–1797, 2017.

[16] D. Angley, B. Ristic, S. Suvorova, B. Moran, F. Fletcher, H. Gaetjens, and S. Simakov, "Non-myopic sensor scheduling for multistatic sonobuoy fields," *IET Radar, Sonar & Navigation*, vol. 11, no. 12, pp. 1770–1775, 2017.

[17] D. Angley, S. Suvorova, B. Ristic, W. Moran, F. Fletcher, H. Gaetjens, and S. Simakov, "Sensor scheduling for target tracking in large multistatic sonobuoy fields," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 3146–3150.

[18] B.-T. Vo, C.-M. See, N. Ma, and W. T. Ng, "Multi-sensor joint detection and tracking with the Bernoulli filter," *IEEE Trans. Aerospace and Electronic Systems*, 2012.

[19] B. Ristic and A. Farina, "Target tracking via multi-static doppler shifts," *IET Radar, Sonar & Navigation*, vol. 7, no. 5, pp. 508–516, 2013.

[20] D. Musicki and B. La Scala, "Multi-target tracking in clutter without measurement assignment," *IEEE Trans. Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 877–896, July 2008.

INTERFERENCE SUPPRESSION USING ADAPTIVE NULLING ALGORITHM WITHOUT CALIBRATION SOURCES

Peng Chen† *Wei Wang* *Jingjie Gao*

School of Information Engineering, Chang'an University, Xi'an, 710064, China

ABSTRACT

Interference suppression using adaptive nulling algorithm is an important array signal processing technique for radar/sonar sensing. However, in long term task, most of the arrays' parameters vary from time to time, which need known sources to re-calibrate. To be free of calibration sources, this paper presents an adaptive nulling algorithm using array observation data. We first establish the model of steering vector (SV) mismatches due to gain-phase error and sensor shifting. Then the angle-related bases of received signal subspace are estimated by applying a joint optimization method consists of Genetic algorithm (GA) and quasi-Newton method. In the end, the array weighting vector can be calculated, and the results of several numerical simulations are demonstrated, which shows that the proposed algorithm can significantly improve the interference suppression performance of sensor array.

Index Terms— Adaptive nulling, Signal subspace, Steering vector estimation,Uncalibrated array

1. INTRODUCTION

Using autonomous systems-based sensor arrays to detect and locate targets is a research direction that has attracted much attention in recent years. Currently, sensor arrays are usually calibrated before the autonomous system is put into use. However, the autonomous system may face environmental changes such as temperature and pressure during the task, and the sensors carried by it may have errors in amplitude, phase, and array shape. These errors are generally unpredictable before the autonomous system is put into use. In actual applications, it may cause a serious decrease in the interference suppression performance of the sensor array.

During the past 30 years, plenty of adaptive nulling methods had been proposed to decrease the effects of calibration error. By adding a scaled identity matrix to the sample covariance matrix (SCM), Cox proposed the diagonal loading beamformer [1]. Because of the precise selection of the diagonal level is hard to obtain, researches proposed robust Capon beamformer [2][3] and the worst-case optimization beamformer [4]. Based on this, convex optimization [5], linear constrains [6] and the iterative approaches [7] were applied to enhance the robustness of beamforming. These method mainly focused on maintaining main beam towards the desired direction in the case of calibration errors. However, when the SV mismatch is severe, these beamformers may suffer from performance degradation, especially in a high signal-to-noise ratio (SNR) environment [8].

Plenty of works show that the singal of interest (SOI) component in the SCM is the main cause of beamformer's performance degradation [8]. To create a covariance matrix that is free of the SOI component, an angular-sector-based covariance matrix reconstruction and estimation beamformer (REB) was proposed in [9]. Several categories of interference-plus-noise covariance matrix (INCM)-based beamformers were then proposed, such as the sparse reconstruction methods [10], and the low-complexity methods [11][12], subspace-based method [13]. In [14], an annulus-uncertainty-set-based method was proposed to alleviate the performance degradation due to random SV mismatch. In[15], the proposed algorithm corrected all SVs of possible interferences, and then corresponding interference power were estimated , which achieved satisfactory performance with high computational complexity. [10] demonstrated an INCM reconstruction-based adaptive beamformer for co-prime array, which shows effectiveness in suppressing interference. Recently, weighted subspace fitting-based methods were proposed to overcome sensor position error [16][17]. However, when other kind of error exist, these methods may not correctly fit the signal subspace, which may cause failure in interference suppression.

In this paper, we propose an adaptive nulling algorithm that only requires the number of signals. By modeling the SV mismatches due to various causes, we establish a joint optimization problem using the idea of signal subspace fitting. Next, instead of estimating powers of all signals, we reconstruct the INCM by extending the subspace bases transition to eigen-space bases transition. In the end, SV of the SOI is estimated and the results of simulations validate the performance of the proposed algorithm.

This work is supported by the China Postdoctoral Science Foundation under Grant 2019M660049XB, by the Fundamental Research Funds for the Central Universities, CHD under Grant 300102240302, and by the National Natural Science Foundation of China under Grants 61871059 and 61901057.

2. PROBLEM FORMULATION

Consider an array with *M* omnidirectional sensors that receives far-field narrowband signals from several sources. The array observation data at the k-th snapshot can be written as

$$\mathbf{x}(k) = \mathbf{a}_0 s_0(k) + \sum\nolimits_1^Q \mathbf{a}_q s_q(k) + \mathbf{n}(k) \,, \tag{1}$$

where $\mathbf{a}_0$ and $\mathbf{a}_q$ denote the actual SVs of the desired signal and the q-th interference, respectively. s_0, s_q, and $\mathbf{n}(k)$ denote the waveform of the desired signal, the q-th interference, and the additive white Gaussian noise vector, respectively. We assume that the desired signal, interferences, and noise to be uncorrelated with each another. To overcome the sensor displacement and gain-phase error, we start from modeling the mismatched SV from the nominal SV as

$$\begin{aligned} \mathbf{a}(\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \theta) &= \alpha \odot e^{j[k_\mathrm{w}(\bar{\mathbf{d}}+\mathbf{d}_\mathrm{e})\sin\theta+\varphi_\mathrm{e}]} \\ &= \bar{\mathbf{a}}(\theta) \odot \alpha \odot e^{j(k_\mathrm{w}\mathbf{d}_\mathrm{e}\sin\theta+\varphi_\mathrm{e})} \,, \end{aligned} \tag{2}$$

where $\bar{\mathbf{d}}$ is the assumed sensor position vector, $\mathbf{d}_\mathrm{e}$ denotes the sensor position error vector, $\odot$ denotes the Hadamard product, and k_w is the wavenumber. α and φ_e are the angle-independent sensor gain vector and phase error vector, respectively. Usually, the first sensor is the reference sensor in the array, which is assumed without sensor position error and phase error, and the sensor gain for the first sensor is 1. Therefore, $\mathbf{d}_\mathrm{e}$, α and φ_e can be expressed as

$$\begin{aligned} \mathbf{d}_\mathrm{e} &= [0, d_2, \cdots, d_M]^\mathrm{T} \in \mathbb{R}^{M\times 1} \\ \alpha &= [1, \alpha_2, \cdots, \alpha_M]^\mathrm{T} \in \mathbb{R}^{M\times 1} \\ \varphi_\mathrm{e} &= [0, \varphi_2, \cdots, \varphi_M]^\mathrm{T} \in \mathbb{R}^{M\times 1} \,. \end{aligned} \tag{3}$$

The SCM contains the information about the actual array calibration and the signals, we can eigen-decompose the SCM $\hat{\mathbf{R}}_\mathrm{x}$ as

$$\begin{aligned} \hat{\mathbf{R}}_\mathrm{x} &= \sum_{m=1}^{M} \lambda_m \mathbf{v}_m \mathbf{v}_m^\mathrm{H} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\mathrm{H} \\ &= \mathbf{V}_\mathrm{S}\boldsymbol{\Lambda}_\mathrm{S}\mathbf{V}_\mathrm{S}^\mathrm{H} + \mathbf{V}_\mathrm{N}\boldsymbol{\Lambda}_\mathrm{N}\mathbf{V}_\mathrm{N}^\mathrm{H} \,, \end{aligned} \tag{4}$$

where λ_m and $\mathbf{v}_m$ are the m-th eigenvalue in descending order and the corresponding eigenvector, respectively. $\boldsymbol{\Lambda} = diag\{\lambda_1, \lambda_2, \cdots, \lambda_M\}$ is a diagonal matrix that consists of all eigenvalues in a descending order, $\mathbf{V}$ is the matrix that contains all eigenvectors. $\boldsymbol{\Lambda}_\mathrm{S} = diag\{\lambda_1, \lambda_2, \cdots, \lambda_L\}$ contains L dominant eigenvalues, and $\mathbf{V}_\mathrm{S}$ is the signal subspace that contains the corresponding eigenvectors. $\boldsymbol{\Lambda}_\mathrm{N} = diag\{\lambda_{L+1}, \cdots, \lambda_M\}$ consists of the remaining eigenvalues, and $\mathbf{V}_\mathrm{N}$ denotes the noise subspace that contains the corresponding eigenvectors.

3. PROPOSED ALGORITHM

In this section, we establish a hybrid optimization problem to estimate the angle-related bases consist of signal SVs using subspace fitting technique. Then we propose a novel INCM reconstruction method that directly eliminate the desired signal component from the sample covariance matrix.

3.1. Angle-related bases estimation

When the precise information about the array and the signals are exactly known, the signal subspace equals to the space spanned by the actual SVs of signals, which is

$$\mathrm{span}\{\mathbf{V}_\mathrm{S}\} = \mathrm{span}\{\mathbf{A}\} \,, \tag{5}$$

where $\mathbf{A} = [\mathbf{a}(\theta_0), \mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_0)]$ denote the actual SV set consists of all Q+1 signal SVs. It is worth noticing that when the interference is coherent with desired signal, or the INR is extremely higher than the SNR, the number of dominate eigenvalues L does not equal the number of signals. However, the number of signals can be estimated using various algorithms.

According to (2), the gain errors is independent of phase errors, then the gain error of the m-th sensor can be estimated as

$$\hat{\alpha}_m = \sqrt{\frac{\hat{\mathbf{R}}_\mathrm{x}(m,m) - \lambda_M}{\hat{\mathbf{R}}_\mathrm{x}(1,1) - \lambda_M}} \,, \tag{6}$$

where $\hat{\mathbf{R}}_\mathrm{x}(m,m)$ denotes the m-th diagonal elements of the SCM, and λ_M is the smallest eigenvalue of the SCM. It can be seen that the sensor position error and the phase error only influence the phase of the SV. It is difficult to accurately estimate the precise directions, sensor position errors and the phase errors separately because $k_\mathrm{w}\mathbf{d}_\mathrm{e}\sin\theta$ in (2) can be treated as angle-related phase errors, which is coupled with the angle-independent phase errors φ_e. However, we can estimate the mismatched SV set by minimizing the difference of the signal subspace and the space spanned by the possible mismatched SV set as

$$\hat{\mathbf{A}}(\hat{\mathbf{d}}_\mathrm{e}, \hat{\varphi}_\mathrm{e}, \hat{\Theta}) = \min_{\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \Theta} \mathrm{tr}\{\, \mathbf{P}^\perp \mathbf{V}_\mathrm{S} \mathbf{W} \mathbf{V}_\mathrm{S}^\mathrm{H}\} \,, \tag{7}$$

where $\mathbf{W}$ is a positive definite weighting matrix, which equals to $(\boldsymbol{\Lambda}_\mathrm{S} - \lambda_M \mathbf{I})^2 \boldsymbol{\Lambda}_\mathrm{S}^{-1}$ under the condition of lowest asymptotic variance and $\mathbf{P}^\perp$ is the orthogonal projection matrix, which is formed as $\mathbf{P}^\perp = \mathbf{I} - \mathbf{A}(\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \Theta)\mathbf{A}^+(\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \Theta)$, where $\mathbf{A}(\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \Theta) \in \mathbb{C}^{M\times(Q+1)}$ represents the possible mismatched SV set, $(\cdot)^+$ denotes the Moore-Penrose inversion, and $\mathbb{C}$ is the complex number field. Θ consist of possible directions of all signals. The i-th column in $\mathbf{A}(\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \Theta)$ represents the possible SV of the i-th signal, which can be formed as

$$\mathbf{a}(\mathbf{d}_\mathrm{e}, \varphi_\mathrm{e}, \theta_i) = \hat{\alpha} \odot e^{j[k_\mathrm{w}(\mathbf{d}+\mathbf{d}_\mathrm{e})\sin\theta_i+\varphi_\mathrm{e}]} \,, \tag{8}$$

where $\hat{\alpha} = [1, \hat{\alpha}_2, \cdots, \hat{\alpha}_2]^\mathrm{T}$ is the estimated gain error vectors in (5). The minimization problem in (6) is obvious a nonlinear optimization problem with $2M + Q - 1$ variables. The

previous work in [17] use GA to tackle this problem. However, with large number of variables, the GA requires large number of generations or iterations to present a satisfactory result. Therefore, we use a joint optimization method that initialize with a few generations of GA, then we use a quasi-Newton Method called the BFGS method to tackle this optimization problem. First, we need to construct the solution vector of the minimization problem as

$$\delta = [d_2, \cdots, d_M, \varphi_2, \cdots, \varphi_M, \theta'_0, \cdots, \theta'_Q]^{\mathrm{T}} \tag{9}$$

where $\theta'_q, q = 0, 1, \cdots, Q$ is the possible DOAs of all signal in an ascending order, and θ'_0 is not necessarily the DOA of the SOI. (6) can be rewritten as

$$\hat{\mathbf{A}}(\hat{\delta}) = \min_{\delta} F(\delta) \tag{10}$$

where $F(\delta)$ is the objective function in (6). Hence, the iteration algorithm of the BFGS method can be formed as

$$\hat{\delta}_{(l+1)} = \hat{\delta}_{(l)} - \beta_{(l)}[F''(\hat{\delta}_{(l)})]^{-1}F'(\hat{\delta}_{(l)}) \tag{11}$$

where $\hat{\delta}_{(l)}$ and $\beta_{(l)}$ are the solution vector and step length at lth iteration, respectively. $F'(\delta)$ and $F''(\delta)$ indicate the gradient and Hessian of $F(\delta)$, respectively. and the gradient can be obtained as

$$F'(\delta) = [\frac{\partial F}{\partial d_m}, \cdots, \frac{\partial F}{\partial \varphi_m}, \cdots, \frac{\partial F}{\partial \theta'_0}, \cdots, \frac{\partial F}{\partial \theta'_Q}]^{\mathrm{T}}, \tag{12}$$

where the $\partial F/\partial d_m$ denotes the partial derivative of variable d_m. The close-form of partial derivative is difficult to obtain, we can approximate the partial derivative by central difference. For example, $\partial F/\partial d_2$ can be approximated as

$$\frac{F(\hat{\delta}_{(l)})}{\partial d_2} \approx \frac{F(\hat{\delta}_{(l)} + \Delta\delta_{d_2}) - F(\hat{\delta}_{(l)} - \Delta\delta_{d_2})}{2\Delta d_2}, \tag{13}$$

where $\Delta\delta_{d_2} = [\Delta d_2, \mathbf{0}]^{\mathrm{T}}$, and Δd_2 denotes a very small positive value. By using the central difference method, the gradient of $F(\delta)$ can be efficiently calculated. Moreover, the close-form of Hessian matrix is difficult to obtain. By utilizing the BFGS method, the inversion of the Hessian matrix $[F''(\hat{\delta}_{(l)})]^{-1}$ can be obtained. However, the BFGS method is sensitive to the initial value $\hat{\delta}_0$, when $\hat{\delta}_0$ is far from the real values, the BFGS may fail to converge. Considering the sensitivity of BFGS, we can estimate initial values of BFGS by using a global optimization method such as the GA with small number of generations. By combining the GA and the BFGS method, the mismatched SV set in (6) can be estimated as

$$\hat{\mathbf{A}}_{\mathrm{S}} = \hat{\mathbf{A}}(\hat{\mathbf{d}}_{\mathrm{e}}, \hat{\varphi}_{\mathrm{e}}, \hat{\Theta}) = F(\hat{\delta}). \tag{14}$$

It is worth noticing that though the difference between span$\{\hat{\mathbf{A}}_{\mathrm{S}}\}$ and span$\{\mathbf{V}_{\mathrm{S}}\}$ is minimized, the estimated parameters $\hat{\mathbf{d}}_{\mathrm{e}}$, $\hat{\varphi}_{\mathrm{e}}$, $\hat{\Theta}$ is not necessarily accurate because these parameters are coupled together and cannot be precisely and separately estimated in this method.

3.2. Covariance matrix reconstruction

To avoid using the estimated power of interference and noise, we can reconstruct the INCM by eliminating the SOI component directly from the SCM using subspace techniques. With well estimated mismatched SV set $\hat{\mathbf{A}}_{\mathrm{S}}$, (4) can be rewritten as span$\{\hat{\mathbf{A}}_{\mathrm{S}}\} \approx$ span$\{\mathbf{V}_{\mathrm{S}}\}$, where $\mathbf{V}_{\mathrm{S}}$ can be seen as a set of orthogonal bases of the signal subspace, and $\hat{\mathbf{A}}_{\mathrm{S}}$ can be regarded as a set of angle-related non-orthogonal bases of the signal subspace. Because the signal subspace is orthogonal to the noise subspace, each column in $\hat{\mathbf{A}}_{\mathrm{S}}$ is orthogonal to the noise subspace, which can be expressed as span$\{\hat{\mathbf{A}}_{\mathrm{S}}\}\perp$span$\{\mathbf{V}_{\mathrm{N}}\}$.

Therefore, span$\{\hat{\mathbf{A}}_{\mathrm{S}}\}$ is the orthogonal complement of span$\{\mathbf{V}_{\mathrm{N}}\}$ in span$\{\mathbf{V}\}$, which indicates that

$$\text{span}\{[\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]\} \approx \text{span}\{[\mathbf{V}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]\} = \text{span}\{\ \mathbf{V}\}, \tag{15}$$

where $[\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]$ is an $M \times M$ matrix, and each column in $[\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]$ can be seen as a basis of the observation space span$\{\mathbf{V}\}$. Hence, we define a bases transition matrix from $[\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]$ to $\mathbf{V}$, and $\mathbf{V}$ is

$$\mathbf{T} = [\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]^{+}\mathbf{V},\ \mathbf{V} = [\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]\mathbf{T}. \tag{16}$$

Then the SCM in (4) can be rewritten as

$$\hat{\mathbf{R}}_{\mathrm{x}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{H}} = [\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]\mathbf{T}\mathbf{\Lambda}\mathbf{T}^{\mathrm{H}}[\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]^{\mathrm{H}}. \tag{17}$$

To eliminate the component of the SOI from the SCM, the INCM can be directly reconstructed as

$$\hat{\mathbf{R}}_{\mathrm{i+n}} = [\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]\mathbf{D}\mathbf{T}\mathbf{\Lambda}\mathbf{T}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}[\hat{\mathbf{A}}_{\mathrm{S}}\ \mathbf{V}_{\mathrm{N}}]^{\mathrm{H}}, \tag{18}$$

where $\mathbf{D}$ denotes a $M \times M$ diagonal matrix as

$$\mathbf{D} = diag\{[\mu, \mathbf{1}_{1\times M-1}]\}. \tag{19}$$

where μ is a very small positive constant that ensure the SOI component can be eliminated. $\mu = 0$ is not recommended because it may result in zero eigenvalue or extremely small positive eigenvalues of $\hat{\mathbf{R}}_{\mathrm{i+n}}$, which may cause the INCM noninvertible.

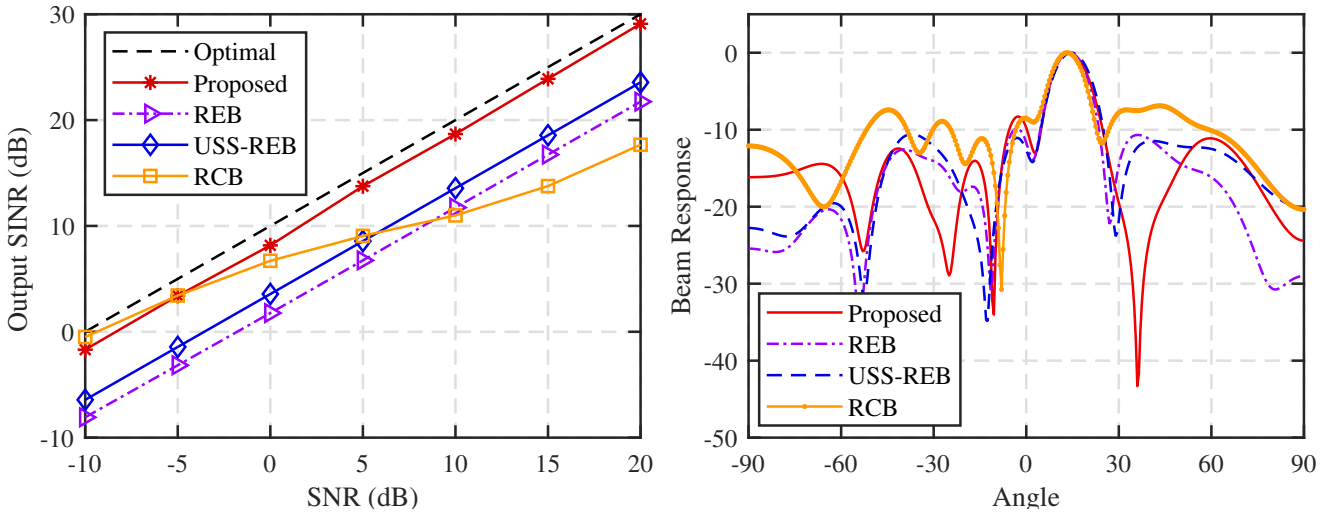
3.3. SOI SV estimation

The actual SV of the SOI is usually unavailable in practical applications, which needs to be estimated or corrected. In this subsection, the desire signal covariance matrix (DSCM) will be reconstructed using the idea similar to the INCM reconstruction, and the SV of the SOI can be estimated from the reconstructed DSCM.

$$\hat{\mathbf{R}}_{\mathrm{s}}(f_{\mathrm{r}}) = \int_{\Theta_{\mathrm{s}}} \frac{\mathbf{a}(\theta, f_{\mathrm{r}})\mathbf{a}^{\mathrm{H}}(\theta, f_{\mathrm{r}})}{\mathbf{a}^{\mathrm{H}}(\theta, f_{\mathrm{r}})\hat{\mathbf{R}}_{\mathrm{F}}^{-1}\mathbf{a}(\theta, f_{\mathrm{r}})} d\theta, \tag{20}$$

where Θ_{s} is the angular sector of the SOI. Unlike $\hat{\mathbf{R}}_{\mathrm{i+n}}$, $\hat{\mathbf{R}}_{\mathrm{s}}$ is supposed to contain only the SOI that originates from $\theta_0 \in \Theta_{\mathrm{s}}$, and the SV of the SOI can be estimated as

$$\hat{\mathbf{a}}_0 = \sqrt{M}\mathcal{P}\{\hat{\mathbf{R}}_{\mathrm{x}} - \hat{\mathbf{R}}_{\mathrm{i+n}}\} \tag{21}$$



(a) Output SINR versus input SNR.

(b) Comparision of normalized beampattern.

Fig. 1: Performance comparison of different adaptive beamformers.

where $\mathcal{P}\{\cdot\}$ denotes the eigenvector corresponding to the largest eigenvalue of a Hermitian matrix. Therefore, the weighting vector of the proposed focused wideband beamformer can be written as

$$\mathbf{w} = \frac{\hat{\mathbf{R}}_{i+n}^{-1}\hat{\mathbf{a}}_0}{\mathbf{a}_0^H\hat{\mathbf{R}}_{i+n}^{-1}\hat{\mathbf{a}}_0} \tag{22}$$

4. NUMERICAL SIMULATION

In this section, we consider an non-ideal scenario to evaluate the robustness of the proposed beamformer. We assume that a uniform linear array with 10 omnidirectional sensors receive signals from three far-field sources. The sensors are assumed evenly spaced at half wavelength. Two interferences with interference-to-noise ratio at 20 dB impinge from $-25°$ and $35°$, and the desired signal impinges from $15°$. The desired signal and interferences are generated from zero means complex Gaussian noises and therefore spatially and temporally independent. In the case of SINR versus SNR, the number of snapshots is fixed at $K = 30$, and in the case of SINR versus snapshots number, the SNR is fixed at 10 dB.

The results are averages of 200 Monte-Carlo simulations. In these simulations, We assumed that the calibration error is partially caused by gain and phase perturbations in each sensor, which distributed in $\mathcal{N}(0, 0.1^2)$ and $\mathcal{N}(0, (0.1\pi)^2)$. Besides, the calibration error contains sensor position error, which is a normal distribution in $\mathcal{N}(0, 0.1^2)$ except for the reference sensor. The proposed beamformer is compared with 3 beamformers, namely the RCB [2], REB [9] and USS-REB [18]. All tested beamformers are compared in the scale of output SINR.

Fig. 1a shows output SINR of four different beamformer. In this case, signal direction error is subject to uniform distribution in $[-2°, 2°]$. With the increase of SNR, the performance of RCB degrades severely due to the SOI component in the SCM. For REB and USS-REB, the output SINRs are lower than RCB when the SNR is smaller than 5 dB. Although proposed beamformer performs worse than RCB when the SNR is -10 dB, it can avoid self-cancellation at high SNR and efficiently suppress interferences in the case of multiple calibration error.

Fig. 1b demonstrate the normalized beam-pattern when SNR = 20 dB. It is obvious that all tested beamformer can steer the main lobe to $\theta_0 = 15°$. However, RCB fails to form nulls around the actual direction of two interferences, the nulls of REB and USS-REB deviate from actual direction of interferences. The proposed beamformer can form two nulls precisely at $\theta_0 = -25°$ and $\theta_0 = 35°$. In other words, the proposed beamformer can suppress interferences in the case of multiple calibration error without known calibration sources.

5. CONCLUSION

This paper proposes an adaptive beamforming and nulling method, which aims to suppress interferences based on the array observation data, and calculate the beamformer's weighting vector using covariance matrix reconstruction. A set of angle-related bases of signal subspace are estimated by applying a joint optimization method. By applying eigen-space bases transition method, the INCM is reconstructed by eliminating SOI component in the SCM, and then the SV of SOI is estimated. The results of several numerical simulations show the performance of proposed beamformer.

6. REFERENCES

[1] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 35, no. 10, pp. 1365–1376, 1987.

[2] P. Stoica, Wang Zhisong, and Li Jian, "Robust capon beamforming," *IEEE Signal Processing Letters*, vol. 10, no. 6, pp. 172–175, 2003.

[3] J. Li, P. Stoica, and Z. S. Wang, "Doubly constrained robust capon beamformer," *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2407–2423, 2004.

[4] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 313–324, 2003.

[5] A. Hassanien, S. A. Vorobyov, and K. M. Wong, "Robust adaptive beamforming using sequential quadratic programming: An iterative solution to the mismatch problem," *IEEE Signal Processing Letters*, vol. 15, pp. 733–736, 2008.

[6] S. D. Somasundaram, "Linearly constrained robust capon beamforming," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5845–5856, 2012.

[7] S. E. Nai, W. Ser, Z. L. Yu, and H. W. Chen, "Iterative robust minimum variance beamforming," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1601–1611, 2011.

[8] Rammohan Mallipeddi, Joni Polili Lie, Sirajudeen Gulam Razul, PN Suganthan, and Chong Meng See, "Robust adaptive beamforming based on covariance matrix reconstruction for look direction mismatch," *Progress In Electromagnetics Research Letters*, vol. 25, no. 1, pp. 37–46, 2011.

[9] Y. J. Gu and A. Leshem, "Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3881–3885, 2012.

[10] Chengwei Zhou, Yujie Gu, Shibo He, and Zhiguo Shi, "A robust and efficient algorithm for coprime array adaptive beamforming," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1099–1112, 2018.

[11] H. Ruan and R. C. de Lamare, "Robust adaptive beamforming using a low-complexity shrinkage-based mismatch estimation algorithm," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 60–64, 2014.

[12] H. Ruan and R. C. de Lamare, "Robust adaptive beamforming based on low-rank and cross-correlation techniques," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3919–3932, 2016.

[13] Xiaolei Yuan and Lu Gan, "Robust adaptive beamforming via a novel subspace method for interference covariance matrix reconstruction," *Signal Processing*, vol. 130, pp. 233–242, 2017.

[14] L. Huang, J. Zhang, X. Xu, and Z. F. Ye, "Robust adaptive beamforming with a novel interference-plus-noise covariance matrix reconstruction method," *IEEE Transactions on Signal Processing*, vol. 63, no. 7, pp. 1643–1650, 2015.

[15] Zhi Zheng, Yan Zheng, Wen-Qin Wang, and Hongbo Zhang, "Covariance matrix reconstruction with interference steering vector and power estimation for robust adaptive beamforming," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8495–8503, 2018.

[16] P. Chen, Y. Yang, Y. Wang, and Y. Ma, "Robust adaptive beamforming with sensor position errors using weighted subspace fitting-based covariance matrix reconstruction," *Sensors (Basel)*, vol. 18, no. 5, pp. 1476–12, 2018.

[17] P. Chen, Y. Yang, Y. Wang, and Y. Ma, "Adaptive beamforming with sensor position errors using covariance matrix construction based on subspace bases transition," *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 19–23, 2019.

[18] Peng Chen and Yixin Yang, "An uncertainty-set-shrinkage-based covariance matrix reconstruction algorithm for robust adaptive beamforming," *Multidimensional Systems and Signal Processing*, vol. 32, no. 1, pp. 263–279, 2021.

LEARNING ROBUST FEATURES FOR 3D OBJECT POSE ESTIMATION

Christos Papaioannidis and Ioannis Pitas

Department of Informatics, Aristotle University of Thessaloniki, Greece

ABSTRACT

Object pose estimation remains an open and important task for autonomous systems, allowing them to perceive and interact with the surrounding environment. To this end, this paper proposes a 3D object pose estimation method that is suitable for execution on embedded systems. Specifically, a novel multi-task objective function is proposed, in order to train a Convolutional Neural Network (CNN) to extract pose-related features from RGB images, which are subsequently utilized in a Nearest-Neighbor (NN) search-based post-processing step to obtain the final 3D object poses. By utilizing a symmetry-aware term and unit quaternions in the proposed objective function, our method yielded more robust and discriminative features, thus, increasing 3D object pose estimation accuracy when compared to state-of-the-art. In addition, the employed feature extraction network utilizes a lightweight CNN architecture, allowing execution on hardware with limited computational capabilities. Finally, we demonstrate that the proposed method is also able to successfully generalize to previously unseen objects, without the need for extra training.

Index Terms— 3D object pose estimation, Multi-task learning, Convolutional Neural Networks.

1. INTRODUCTION

Autonomous robots or systems are being increasingly employed in several industries (e.g., transportation, construction) to assist in simple or more complex tasks. However, their safe and successful operation in real-world scenarios requires advanced understanding of the surrounding environment. Object pose estimation is critical in this case, as it enables predicting the 3D poses of objects of interest in their surroundings, in order to autonomously take the correct actions according to a given objective. For example, in a human-Unmanned Aerial Vehicle (UAV or drone) collaboration scenario, the UAV should be able to estimate the 3D pose of a tool in order to grab it (e.g., by using a robotic arm) and pass it to a human worker.

Early deep learning-based methods addressed the 3D object pose estimation problem by training a Convolutional Neural Network (CNN) to directly regress [1, 2] or classify [3] an object image to its 3D pose. More recent 6D object pose estimation methods [4, 5, 6, 7, 8] utilized state-of-the-art object detection [9] and instance segmentation [10] methods to first localize the objects of interest in the 2D image (e.g., by regressing their 2D bounding boxes or a set of predefined keypoints) and then computed the final 6D object poses by using the 2D detections in a PnP algorithm. However, these approaches either lack increased 3D object pose estimation accuracy, or rely on very deep neural network architectures, which do not allow fast execution on embedded systems.

In an alternative approach, the final 3D object pose predictions can be obtained indirectly [11, 12, 13, 14, 15, 16]. That is, a lightweight CNN is first utilized to extract object image features, which are then matched with a set of precomputed database entries that represent orientation classes via a Nearest Neighbor (NN) search. While these methods managed to achieve increased 3D object pose estimation performance despite only using a lightweight CNN, non-trivial object symmetries [11, 12, 14, 15, 16] and poor 3D pose representation [13] can cause convergence issues during CNN training.

In this work, we propose a 3D object pose estimation method that aims to overcome all the aforementioned limitations of existing methods. More specifically, we improve the existing feature learning methods by introducing a novel multi-objective loss function to train a lightweight CNN to extract pose-related object image features. The proposed loss function utilizes unit quaternions to represent 3D poses and a symmetry-aware term to handle non-trivial symmetries of objects which facilitate the feature learning process, resulting in discriminative pose-related features from which the final 3D object poses can be more accurately obtained. Moreover, since the proposed loss function is carefully designed to encode 3D object poses in the extracted features, the proposed method is able to successfully generalize to previously unseen objects. Finally, the lightweight architecture of the feature extraction CNN allows execution on hardware with limited computational capabilities, rendering the proposed method suitable for autonomous systems.

In summary, this paper offers the following contributions:

- a novel multi-task objective function for 3D pose feature learning that combines unit quaternions and a symmetry-aware term,

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement numbers 731667 (MULTIDRONE) and 871479 (AERIAL-CORE).

- a 3D object pose estimation method with increased 3D object pose estimation performance and generalization ability that is suitable for embedded systems.

2. PREVIOUS FEATURE LEARNING 3D OBJECT POSE ESTIMATION METHODS

Feature learning pose estimation methods [11, 13, 14, 15, 16, 17] offer several advantages over pose regression [1, 4] or classification [3] approaches, as they only require a shallow CNN architecture, they are scalable to the number of objects [14] and can simultaneously perform object classification. In this case, a CNN is first trained to extract pose-related features from an object image. At test time, the extracted object image features are matched with a set of precalculated database image features via NN search, returning as final object class and 3D pose estimates the ones that correspond to the retrieved closest database image.

In order to learn such pose-related image features, a feature learning pose estimation framework was firstly introduced in [11], where a lightweight CNN was trained to calculate discriminative features using Siamese [18] and triplet [19] network architectures. The feature learning loss function which was used in [11] to train the CNN was of the following form:

$$\mathcal{L} = \mathcal{L}_d + \lambda_w \|\mathbf{w}\|_2^2, \tag{1}$$

where $\mathcal{L}_d$ is a feature learning term, λ_w is a regularization parameter and $\|\mathbf{w}\|_2$ is the L_2-norm of the network parameter vector. $\mathcal{L}_d$ consists of a pairwise and a triplet loss function, $\mathcal{L}_d = \mathcal{L}_{pairs} + \mathcal{L}_{triplets}$ [11], in order to learn object image features from which both the 3D object pose and the object class can be retrieved. $\mathcal{L}_{pairs}$ is responsible for keeping object images with similar poses close in the feature space, while $\mathcal{L}_{triplets}$ aims to distinguish between different object identities. Later, the total loss function (1) was extended in [14] by adding a dynamic margin in the triplet loss function $\mathcal{L}_{triplets}$ to improve the robustness of the resulting low-dimensional features. The dynamic margin, which utilized unit quaternions and quaternion distance, was defined as:

$$\varepsilon_d = \begin{cases} 2\arccos(|\mathbf{q}_i^T \mathbf{q}_j|) & if\ c_i = c_j, \\ n & \text{else, for } n > \pi, \end{cases} \tag{2}$$

where $\mathbf{q}_i$, $\mathbf{q}_j$ and c_i, c_j are the corresponding ground truth 3D pose and object class labels of training samples s_i, s_j, respectively, while n is a constant value for penalizing pairs from different object classes.

In order to learn more discriminative features for 3D object pose estimation and object recognition, a pose-guided pairwise loss and an extra 3D pose regression term was used in the total loss function in [13]:

$$\mathcal{L} = \mathcal{L}_{pose} + \mathcal{L}_{object} + \mathcal{L}_{reg} + \lambda \|\mathbf{w}\|_2^2. \tag{3}$$

In this case, the pairwise loss $\mathcal{L}_{pose}$ enforces a direct relationship between the learned features and the real pose label differences, while the role of triplet loss $\mathcal{L}_{object}$ remains the same as in [11, 14]. The trained model yielded pose-related features that greatly improved the 3D object pose estimation performance compared to [11]. The work of [14] was also extended in [15, 16], where a quaternion regression term was used in the total loss function, along with the feature learning term $\mathcal{L}_d$. By using quaternions and quaternion regression, the trained model not only demonstrated increased 3D object pose estimation performance, but also enabled direct 3D object pose regression by completely omitting the NN search step.

3. MULTI-TASK FEATURE LEARNING

The objective function used in the proposed 3D object pose estimation method is:

$$\mathcal{L} = \lambda_p \mathcal{L}_{pose} + \lambda_o \mathcal{L}_{obj} + \lambda_r \mathcal{L}_{qreg}, \tag{4}$$

where $\mathcal{L}_{pose}$, $\mathcal{L}_{obj}$, $\mathcal{L}_{qreg}$ are the pairwise, triplet and quaternion regression loss functions, respectively, and λ_p, λ_o, λ_r are hyper-parameters used to control the contribution of each term in the total loss function. The key difference between the proposed loss function and (3) is that the terms $\mathcal{L}_{pose}$, $\mathcal{L}_{qreg}$ in (4) utilize unit quaternions and quaternion distance. Unit quaternions $\mathbf{q} \in \mathbb{R}^4$, $\mathbf{q} = [q_0, q_1, q_2, q_3]^T$, $\|\mathbf{q}\|_2 = 1$, offer a preferable alternative for rotations representation, as they are more compact compared to rotation matrices and also avoid the gimbal lock problem [20] of the Euler angle representation. Since unit quaternions double-cover the $\mathcal{SO}(3)$ ($\mathbf{q}$ and $-\mathbf{q}$ represent the same rotation), in the proposed method we enforce $q_0 \geq 0$ in order to achieve a one-to-one correspondence between rotation matrices and quaternions. Moreover, the quaternion distance offer an accurate representation of real pose differences, which is required in the proposed method.

The pairwise ($\mathcal{L}_{pose}$) and the triplet ($\mathcal{L}_{obj}$) loss functions utilized in (4) are specifically designed in order to learn a discriminative feature space during CNN training. Let $s_i = \{\mathbf{x}_i, c_i, \mathbf{q}_i\}, i = 1, \ldots, N$ be a training set sample which contains an RGB-D image $\mathbf{x}_i$ of an object, its assigned object class label $c_i \in \mathcal{C} = \{c_1, \ldots, c_L\}$ and the corresponding 3D pose quaternion $\mathbf{q}_i \in \mathbb{R}^4$. Also, let $\mathcal{P} = \{s_i, s_j\}$, $\mathcal{T} = \{s_i, s_j, s_k\}$ be sets containing training sample pairs and triplets, respectively. The pairwise loss $\mathcal{L}_{pose}$ is computed on pairs $\{s_i, s_j\} \in \mathcal{P}$, where the samples s_i, s_j belong in the same object class $c_l, l = 1, \ldots, L$ and is used to enforce pose similarity within the same object class c_l. Note that some objects may seem very similar under different poses, therefore, we need our model to be able to handle these cases where objects are symmetric in a non-trivial way. To this end, we weight the contribution of each sample in $\mathcal{L}_{pose}$ with the corresponding symmetry-aware term $\phi(\mathbf{q}_i, \mathbf{q}_j)$ to impose infor-

mation about object symmetries to the model. Therefore, if $\mathbf{f}_i = f(\mathbf{x}_i) \in \mathcal{F} \subset \mathbb{R}^d$ are the features obtained from the last fully connected CNN layer having $\mathbf{x}_i$ as input, the pairwise loss is defined as:

$$\mathcal{L}_{pose} = \sum_{s_i, s_j} \phi(\mathbf{q}_i, \mathbf{q}_j) \cdot \{\|\mathbf{f}_i - \mathbf{f}_j\|_2^2 - 2\arccos(|\mathbf{q}_i^T \mathbf{q}_j|)\}^2, \tag{5}$$

where $\phi(\mathbf{q}_i, \mathbf{q}_j) = \|\mathbf{d}_{\mathbf{q}_i} - \mathbf{d}_{\mathbf{q}_j}\|_2^2$ and $\mathbf{d}_{\mathbf{q}_i}$, $\mathbf{d}_{\mathbf{q}_j}$ are the rendered depth object images under the poses $\mathbf{q}_i$, $\mathbf{q}_j$, respectively. Essentially, $\mathcal{L}_{pose}$ forces the Euclidean feature distance between two samples from the same object class to be equal to the quaternion distance between the corresponding 3D poses $\mathbf{q}_i$, $\mathbf{q}_j$. However, as in some cases an object may appear the same when seen from different viewpoints, convergence problems may occur during training. By using $\phi(\mathbf{q}_i, \mathbf{q}_j)$ in (5), the learned features of symmetric samples with very similar rendered depth images $\mathbf{d}_{\mathbf{q}_i}$, $\mathbf{d}_{\mathbf{q}_j}$ are not directly affected by the magnitude of the symmetry-agnostic $2\arccos(|\mathbf{q}_i^T \mathbf{q}_j|)$ term. Therefore, minima closer to the global minimum can be located during the network optimization process.

The triplet loss term $\mathcal{L}_{obj}$ in (4) enforces features coming from same-object class samples to have smaller distances in the feature space, when compared to the distances of features calculated from different object class samples. For this purpose, the sample triplets $\{s_i, s_j, s_k\} \in \mathcal{T}$, consist of samples s_i, s_j coming from the same object class $c_l, l = 1, \ldots, L$, while s_k is a sample coming from any different object class. $\mathcal{L}_{obj}$ is of the following form:

$$\mathcal{L}_{obj} = \sum_{s_i, s_j, s_k} \frac{\|\mathbf{f}_i - \mathbf{f}_j\|_2}{\|\mathbf{f}_i - \mathbf{f}_k\|_2 + \varepsilon}, \tag{6}$$

so that the distance in the feature space between the same object class is forced to be smaller than the distance between object features coming from different classes. ε is a small regularizing constant, that also prevents having a zero denominator in (6).

The quaternion regression term $\mathcal{L}_{qreg}$ is defined as:

$$\mathcal{L}_{qreg} = 2arccos(|\mathbf{q}^T \hat{\mathbf{q}}|), \tag{7}$$

where $\mathbf{q}$, $\hat{\mathbf{q}}$ are the ground truth and the predicted 3D object pose quaternion, respectively. $\mathcal{L}_{qreg}$ not only enables the CNN to directly regress the 3D object pose in the form of a unit quaternion, but also assists the feature learning process by imposing extra information about 3D object poses to the model. However, quaternion regression requires special attention, as the four quaternion entries q_0, q_1, q_2, q_3 are not independent. The term $\sin\frac{\theta}{2}$ is found in all three entries q_1, q_2, q_3, while $\cos\frac{\theta}{2}$ contributes to q_0. Therefore, trying to directly regress $\mathbf{q}$ leads to inferior performance [21]. In contrast, in the proposed method, unit quaternions are obtained implicitly. That is, the independent axis-angle rotation representation entries $\mathbf{r} = [\theta', u_1, u_2, u_3]^T$, $\theta' = \frac{\theta}{2}$ are regressed, where $\mathbf{u} \in \mathbb{R}^3$, $\mathbf{u} = [u_x, u_y, u_z]^T$ is the unit rotation axis and $\theta \in \mathbb{R}$ its rotation angle. Ultimately, the predicted 3D object pose quaternion $\hat{\mathbf{q}} = [\hat{q}_0, \hat{q}_1, \hat{q}_2, \hat{q}_3]$ used in (7) is obtained as follows:

$$\begin{cases} \hat{q}_0 = \cos(\theta') \\ \hat{q}_1 = u_1 \sin(\theta') \\ \hat{q}_2 = u_2 \sin(\theta') \\ \hat{q}_3 = u_3 \sin(\theta'). \end{cases} \tag{8}$$

4. EXPERIMENTAL EVALUATION

The CNN used in the proposed method has the same architecture as the ones used in [11, 13, 14, 15]. The first two network layers are convolutional ones with rectified linear (ReLU) activation function, followed by two fully connected layers. The final fully connected layer produces the 3D pose feature vector $\mathbf{f} \in \mathcal{F} \subset \mathbb{R}^d$. In all the experiments, the feature dimensionality was set to $d = 32$. For the quaternion regression, an extra fully connected layer is added after the feature layer. This layer is followed by the quaternion activation layer, which outputs $\hat{\mathbf{q}}$, according to (8). We empirically set $\lambda_p = 10$, $\lambda_o = 1$ and $\lambda_r = 0.5$ in all our experiments to benefit 3D pose feature learning. The overall CNN is trained for 400 epochs using the stochastic gradient decent method, with momentum 0.9 and initial learning rate of 0.01, which is reduced in each epoch.

The proposed method is compared to the baseline methods of [11, 13, 14, 15]. In all experiments all models were trained using the Cropped LineMOD dataset [11]. It has to be noted that in all cases, the estimated 3D object pose is the ground truth pose assigned to the closest database sample retrieved by the nearest neighbor search. For the quantitative evaluation of all trained models, the angular error metric is used [14]:

$$err(\mathbf{q}, \hat{\mathbf{q}}) = 2\arccos(|\mathbf{q}^T \hat{\mathbf{q}}|), \tag{9}$$

where $\mathbf{q}$, $\hat{\mathbf{q}}$ are the the ground truth and the estimated object pose, respectively. The 3D pose estimation accuracy at threshold t is then defined as the percentage of test samples, for which the angular error between the estimated and the ground truth pose is below a threshold angle t, $err(\mathbf{q}, \hat{\mathbf{q}}) < t$. Note that, the pose estimation accuracy is calculated only for the test samples that were correctly matched to their corresponding object class.

The comparison between the performance of the proposed and the baseline CNN models $3DPOD$ [11], $PEDM$ [14], $PGFL$ [13] and QL [15] for threshold angle values $t \in [5°, 10°, 15°, 20°, 30°, 40°, 45°]$ is presented in Table 1, where the object classification accuracy is also reported. It should be noted that, since the code of [14] could not be made available, the results reported in [14] are directly cited in Table 1 only for threshold angle values $t \in [10°, 20°, 40°]$. As can be seen in Table 1, the proposed method outperforms

Table 1. 3D object pose estimation and object classification accuracy.

	Angular threshold t								
	5°	**10°**	**15°**	**20°**	**30°**	**40°**	**45°**	**Mean (Median) ± Std**	**Object classification**
$3DPOD$ [11]	40.15%	72.72%	86.02%	91.76%	95.42%	96.90%	97.34%	12.75°(7.06°) ± 24.61°	98.94%
$PEDM$ [14] ⋆	-	60.00%	-	93.20%	-	98.00%	-	-	99.30%
$PGFL$ [13]	41.28%	83.07%	93.98%	97.43%	99.11%	99.52%	99.60%	6.89°(5.79°) ± 6.29°	99.64%
QL [15]	41.37%	82.02%	95.32%	98.49%	99.72%	99.92%	99.94%	6.64°(5.78°) ± 5.14°	99.50%
$ours$	**44.13%**	**84.25%**	**95.76%**	**98.77%**	**99.84%**	**99.93%**	**99.94%**	**6.31°(5.53°) ± 4.58°**	**99.68%**

⋆ The results of $PEDM$ are directly cited from [14].

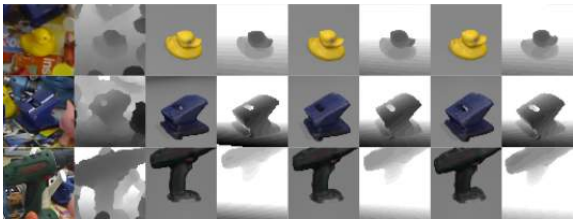


Fig. 1. Retrieved top 3 nearest neighbors for 3 query images from Cropped LineMOD dataset [11]. First two columns show the query RGB-D images and the rest columns depict the retrieved closest nearest neighbors from left to right.

all competing methods, yielding increased 3D object pose estimation accuracy for all angle threshold values. Particularly in the high 3D object pose estimation accuracy area ($t = 5°$) the proposed method increased the 3D object pose estimation accuracy up to 4%. In addition, the proposed method outperforms all competing methods in the object classification task. As also reported in Table 1, the proposed method has lower mean and standard deviation values of the angular error compared to all competing models. The results show that by incorporating the symmetry-aware term ϕ in the quaternion-based feature learning process using (4), the proposed method is able to extract more discriminative pose-related features that enable increased 3D object pose estimation performance.

Apart from the comparison reported in Table 1, we also performed a qualitative evaluation of the proposed method. More specifically, the images of the closest 3 database samples retrieved by the proposed method for random query test images from the Cropped LineMOD dataset are presented in Fig. 1. It can be seen that all query images are successfully matched to database samples that have very similar 3D pose, with the 3D pose difference between them being imperceptible in most cases. Finally, the generalization ability of the proposed method on a previously unseen object (car in random scenes, simulating an objective of a self-driving car scenario) is evaluated by utilizing images from WCVP [22] dataset. Thus, random images from the first split of WCVP dataset were used as query images, while all images from the second split of WCVP dataset were utilized as database samples. Note that, as a pre-processing step, depth images were extracted using the depth estimation method of [23] (since depth images are not provided by WCVP) and then both RGB and depth images were cropped using the ground truth 2D bounding boxes provided by WCVP. The results presented in Fig. 2 show that the proposed method was able to match query images with database samples that have almost identical 3D poses, without the need for an extra training step and despite the fact that cars between query and database images may have different shape or color.

Fig. 2. Generalization ability of the proposed method on previously unseen object. First two columns show the query RGB-D images and the rest columns depict the retrieved closest nearest neighbors from left to right.

5. CONCLUSION

In this work, a 3D object pose estimation method for embedded execution was presented. By utilizing a lightweight CNN and a specifically designed 3D pose feature learning objective function that considers non-trivial object symmetries, the proposed method yielded more discriminating 3D pose features, hence, outperforming state-of-the-art feature learning methods. Experiments in the Cropped LineMOD dataset showed that the proposed method increased the 3D object pose estimation accuracy for all angle threshold values as well as the object classification accuracy. Finally, the proposed method demonstrated increased generalization ability to unseen objects, without the need for extra training.

6. REFERENCES

[1] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox, "Posecnn: A convolutional neural network for 6D object pose estimation in cluttered scenes," *arXiv preprint arXiv:1711.00199*, 2017.

[2] A. Kendall, M. Grimes, and R. Cipolla, "Posenet: A convolutional network for real-time 6-dof camera relocalization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015.

[3] H. Su, C. R. Qi, Y. Li, and L. J. Guibas, "Render for cnn: Viewpoint estimation in images using cnns trained with rendered 3D model views," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015.

[4] T.-T. Do, M. Cai, T. Pham, and I. Reid, "Deep-6Dpose: Recovering 6D object pose from a single rgb image," *arXiv preprint arXiv:1802.10367*, 2018.

[5] M. Rad and V. Lepetit, "Bb8: A scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.

[6] W. Kehl, F. Manhardt, F. Tombari, S. Ilic, and N. Navab, "Ssd-6D: Making rgb-based 3D detection and 6D pose estimation great again," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.

[7] B. Tekin, S. N Sinha, and P. Fua, "Real-time seamless single shot 6D object pose prediction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[8] S. Peng, Y. Liu, Q. Huang, X. Zhou, and H. Bao, "Pvnet: Pixel-wise voting network for 6DoF pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[9] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2016.

[10] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.

[11] P. Wohlhart and V. Lepetit, "Learning descriptors for object recognition and 3D pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[12] W. Kehl, F. Milletari, F. Tombari, S. Ilic, and N. Navab, "Deep learning of local rgb-d patches for 3D object detection and 6D pose estimation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2016.

[13] V. Balntas, A. Doumanoglou, C. Sahin, J. Sock, R. Kouskouridas, and T.-K. Kim, "Pose guided rgbd feature learning for 3D object pose estimation," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.

[14] S. Zakharov, W. Kehl, B. Planche, A. Hutter, and S. Ilic, "3D object instance recognition and pose estimation using triplet loss with dynamic margin," in *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[15] C. Papaioannidis and I. Pitas, "3D object pose estimation using multi-objective quaternion learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2683–2693, 2019.

[16] M. Bui, S. Zakharov, S. Albarqouni, S. Ilic, and N. Navab, "When regression meets manifold learning for object recognition and pose estimation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2018.

[17] C. Papaioannidis, V. Mygdalis, and I. Pitas, "Domain-translated 3D object pose estimation," *IEEE Transactions on Image Processing*, vol. 29, pp. 9279–9291, 2020.

[18] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a siamese time delay neural network," *Advances in Neural Information Processing Systems*, 1994.

[19] E. Hoffer and N. Ailon, "Deep metric learning using triplet network," in *International Workshop on Similarity-Based Pattern Recognition*, 2015.

[20] S. L. Altmann, *Rotations, quaternions, and double groups*, Courier Corporation, 2005.

[21] H.-W. Hsu, T.-Y. Wu, S. Wan, W. H. Wong, and C.-Y. Lee, "Quatnet: Quaternion-based head pose estimation with multi-regression loss," *IEEE Transactions on Multimedia*, vol. 21, no. 4, pp. 1035–1046, 2018.

[22] D. Glasner, M. Galun, S. Alpert, R. Basri, and G. Shakhnarovich, "Viewpoint-aware object detection and continuous pose estimation," *Image and Vision Computing*, vol. 30, no. 12, pp. 923–933, 2012.

[23] S. F. Bhat, I. Alhashim, and P. Wonka, "Adabins: Depth estimation using adaptive bins," *arXiv preprint arXiv:2011.14141*, 2020.

GENERAL FRAMEWORKS FOR ANOMALY DETECTION EXPLAINABILITY: COMPARATIVE STUDY

Ambareesh Ravi, Xiaozhuo Yu, Iara Santelices, Fakhri Karray, Baris Fidan

Dept. of Electrical and Computer Engineering
University of Waterloo, Waterloo, ON, CANADA
{ambareesh.ravi, xz2yu, isanteli, karray, fidan}@uwaterloo.ca

ABSTRACT

Since their inception, AutoEncoders have been very important in representational learning. They have achieved ground-breaking results in the realm of automated unsupervised anomaly detection for various critical applications. However, anomaly detection through AutoEncoders suffers from lack of transparency when it comes to decision making based on the outputs of the AutoEncoder network, especially for image-based models. Though the residual reconstruction error map from the AutoEncoder helps explaining anomalies to a certain extent, it is not a good indicator of the implicitly learnt attributes by the model. A human interpretable explanation of why an instance is anomalous not only enables the experts to fine-tune the model but also establishes and increases trust by non-expert users of the model. Convolutional AutoEncoders in particular suffer the most as there are only limited studies that focus on transparency and explainability. In this paper, aiming to bridge this gap, we explore the feasibility and compare the performances of several State-of-the-Art Explainable Artificial Intelligence (XAI) frameworks on Convolutional AutoEncoders. The paper also aims at providing the basis for future developments of reliable and trustworthy AutoEncoders for visual anomaly detection.

Index Terms— Convolutional AutoEncoders, Explainable AI, Deep Learning

1. INTRODUCTION

The popularity of machine learning has skyrocketed in the last few decades. Deep Learning models in particular are pushing the boundaries of what is possible [1], [2]. With such models becoming increasingly complex in the number of layers and learnable parameters, it is a daunting task to explain the decisions made by these models. Explainable AI (XAI) is a new area of machine learning focusing on making models more interpretable producing explainable models while maintaining performance and enabling the practitioners and users to understand and trust the model. The important traits of a comprehensive explainable model are fairness, trust, reliability, causality and privacy. These traits are especially important when the social, psychological and financial impacts on the predictions are high and also when the domain of operation is not fully fathomed [3]. Currently, the field of XAI is still in its infancy with room for exponential growth.

This work focuses on the usage of some recently developed XAI frameworks on Convolutional AutoEncoders (CAE) for reconstruction-based image anomaly detection tasks. Usually, the *residual reconstruction error map* is used to visualize anomalies. However, the error map can only indicate the absence and partial or deformed reconstruction of anomalous entities in the input. It cannot locate the anomalies precisely and cannot explain why an entity is anomalous based on the learnt notion of normality. Since CAE' decisions are not based on probability scores, most of the existing XAI approaches are incompatible. There is a severe lack of literature on explaining CAE architectures for images and to the best of our knowledge, we are the first to provide a comprehensive and comparative literature survey of XAI studies on CAE design for visual anomaly detection. The main contributions of the paper are (1) a method to reformulate the reconstruction based image anomaly detection models to be used with XAI frameworks that are employed for classification, (2) Comparison of the performances of four popular XAI frameworks based on two different datasets, (3) establishing that adopting state-of-the-art XAI methods to CAEs is more effective than resorting to the residual error maps for human-interpretable visual explanations on anomaly detection.

2. LITERATURE REVIEW

Overviews on various XAI frameworks for convolution networks are discussed in [4], [5]. But existing literature on XAI frameworks developed particularly for AE is very sparse. For non-Convolutional AE two pieces of literature exist [6], [7]. In [6], the work focuses on a custom solution for explaining AE applied to Collaborative Filtering (CF) in recommender systems to predict missing ratings [6]. To explain the recommendation generated from the model an additional explainability vector is introduced. While this approach provides explainability, it is not widely applicable to other networks.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Another work [7] solves the issue of custom explainability model building in [6] through SHapley Additive exPlanations (SHAP) which is shown to be effective in explaining various supervised learning models. The literature leverages kernel SHAP to explain anomalies detected by the AE. The method in [7] explains these anomalies by focusing on the connection between the features with high reconstruction error and the features that are affecting the reconstruction error the most.

The works [6] and [7] mentioned above are applied to fully-connected AEs. In [8], the authors proposed a custom solution to explain the features of CAEs and applied a custom algorithm for satellite images - the CAE is initially trained for reconstruction and used to compute feature-wise distance between bi-temporal images X_1, X_2 using the features from the first hidden layer [8]. The extracted features are compared based on squared error E. A greedy approach is used to select a fraction of the features with the highest standard deviation of the error E [8]. This approach does lead to explainable features but it is not a generalized solution that could be applied to any custom models but to the authors' best knowledge, [8] is the only work applied to CAEs. In this work, we explore the XAI frameworks that can be used for CAEs.

3. METHODS

AutoEncoders (AE) are a family of unsupervised neural network architectures that can efficiently represent data. An AE is constituted by two neural network structures - an *encoder* $E(x)$ that learns to compress the input data x into lower-dimensional latent space representation z and a *decoder* $D(z)$ that learns to decompress the representation z into the original data dimension to reconstruct the input as x'. CAE for anomaly detection in images is trained on a dataset containing normal images and hence can reconstruct the normal images almost perfectly. The inability to reconstruct the unseen anomalous image samples is utilized as a measure of detection i.e., the higher the reconstruction loss between the input and the reconstructed images is the higher the probability of the input being an anomaly is and vice-versa. Reconstruction based approaches using AEs for detecting anomalies do not accompany class predictions. Since most XAI frameworks operate on class predictions and employing a classifier to AE completely changes the paradigm of the solution, we explore four methods that allow us to use the structure of AE without any major structural modifications.

3.1. Layer Relevance Propagation (LRP)

Layer Relevance Propagation (LRP) [9] is based on feature relevance distribution and conservation. LRP uses the trained weights learnt from the forward propagation of inputs to distribute the relevance based on the predictions towards the input features and the relevance values in any given layer is conserved summing up to the same value R across multiple layers and the value R is the prediction score of a particular class c in the output layer whose decision needs to be explained. Equation (1) shows the formula to calculate relevance between neurons in two consecutive layers j and k connected by weight w_{jk} whose activation is given by a. This is the proportion of influence/contribution of neuron in j towards the neuron in layer k. The small positive constant ϵ is added to the denominator that compensates the weakness of relevance propagated to k and it can alleviate the effect of noise, producing sparse explanations.

$$R_j = \sum_k \frac{a_j w_{jk}}{\epsilon + \sum_{0,j} a_j w_{jk}} \tag{1}$$

3.2. Local Interpretable Model-agnostic Explanation (LIME)

Local Interpretable Model-agnostic Explanation (LIME) [10] is a generalized explainability framework that utilizes local *surrogate* models to explain the individual classification or regression decisions of a machine learning model. A surrogate model tends to simplify a complex model by imitating its behaviour to ensure local fidelity. LIME essentially analyses the change in probability scores on multiple instances of a reference input with added noise or change in the value to provide suitable explanations. A set of data samples containing perturbed versions X_p of the reference image x is generated by switching off or replacing the pixels of the interpretable components and the score for each of the samples in the set is calculated. The surrogate model $f_s \epsilon F_s$ is then trained on X_p learning to weight the patches of pixels based on proximity $\pi(x)$ as in equation (2) where f_o is the CAE, F_s is the family of surrogate models, $\Omega(f_s)$ is the complexity of f_s, and $\mathcal{L}$ is the Loss function. Finally, the pixels with the largest weights denote the explanation for the reference image which indicates the essential attribute that makes the model decide on that particular class as follows:

$$explanation(x) = \underset{f_s \epsilon F_s}{\operatorname{argmin}} \mathcal{L}(f_o, f_s, \pi_x) + \Omega(f_s) \tag{2}$$

3.3. SHapley Additive exPlanations (SHAP)

SHapley Additive exPlanations (SHAP) [11] is a feature attribution based explainability method that measures the importance of an input feature concerning the output prediction. SHAP is an additive feature attribution method that uses *Shapely values*, a concept from coalitional game theory that describes how fairly the prediction is distributed among the input features which in our case signifies the quantification of the contribution of each input feature towards the final prediction. Due to this property of shapely values, SHAP provides consistent global interpretation for each data sample. SHAP replaces features with random variables to determine its contribution towards the final output prediction through the

relative difference from the original prediction. The weights for KernelSHAP $\pi_z()$ can be determined by the equation (3) where $|z'|$ is the number of features considered for the coalition and M is the maximum coalition among features.

$$\pi_z(z') = (M-1)/((M/|z'|) \times |z'| \times (M-|z'|)) \quad (3)$$

3.4. Counterfactuals

The counterfactual [12] method of explanation for neural networks is a model-agnostic approach that aims to find the smallest change in feature values that will cause an alteration to the prediction of the model. In anomaly detection, we want to understand the smallest change such that a normal sample is altered as an anomalous sample. Since Counterfactual method is a generalized idea and the algorithm is usually data and model-specific, we apply a custom algorithm to understand the image region that causes the highest reconstruction error (E_r). The algorithm is described in Algorithm 1. Counterfactuals and adversarial examples help in determining the perturbations or changes in input that drive the model towards other extremes of predictions.

Algorithm 1: Greedy Sequential Search

Input: image *I*, anomalous image *I'*, threshold *T*
Output: list of edits $l_E \ni E_r(e, I) > T,\ \forall\ e\ \epsilon\ l_E$
$S \leftarrow Segments(I'),\ l_E \leftarrow [\,],\ currImg = I$
while $E_r < T$ **do**
    let e be the segment that increases E_r *the most*
    $currImg \leftarrow currImg + e$
    Add e to L_e
    Set E_r *as the reconstruction error of* $currImg$
end

4. EXPERIMENTS AND RESULTS

In this section, we present the results of a set of experiments we conducted to compare the performances of the four methods reviewed in Section 3. For our experiments[1], we used a simple convolutional AutoEncoder model with a mirror-identical encoder and decoder. The encoder consists of 5 strided convolutional layers with kernels of size 3×3 and stride 2. The number of kernels in each layer is 64,64,96,96,128 respectively with embedding dimension of 1152×1. Each convolutional layer is followed by *BatchNorm* [13] layer and *ReLU* [14] activation. The decoder is similar to the encoder in reverse with transpose convolutional layers instead and with padding adjustments to maintain output shape. Inputs of size 128x128 are normalized between 0 and 1 without any other pre-processing and the model is trained on batches of size 64 for 300 epochs with a starting learning rate of 1×10^{-3} using Adam optimizer [15] and mean squared error between inputs and reconstructions as objective loss function.

[1]Complete code base containing experiments available at `https://github.com/ambareeshravi/AD_AE_XAI/`

4.1. Modification to AutoEncoders for Applying Generic XAI Frameworks

Most of the XAI algorithms are built around predictions from a classifier and anomaly detection can be considered as a binary classification problem. But reconstruction-based anomaly detection using AutoEncoder do not provide probability scores directly and simple adjustments are made to AutoEncoder architecture for this case. For any dataset, the maximum value of reconstruction error from the anomalous test set (anomaly type) is found and used to normalize the reconstruction losses between 0 and 1. The normalized reconstruction loss can be considered as an anomaly score where a lower score denotes the normality of the input sample. The optimal threshold (δ_o) of anomaly detection (binary classification) is a classification threshold δ_i that can be manually fixed or found using from the set of all n possible thresholds represented by δ in the Receiver Operator Characteristics curve that gives the maximum value of geometric mean of *sensitivity* = true positive rate (TPR) and *specificity* = 1 – False Positive Rate (FPR) as in Equation (4). δ_o can help in finding the required balance between precision and recall according to the application.

$$\delta_o = \underset{\delta_i \epsilon \delta}{\operatorname{argmax}} \sqrt{TPR(\delta_i) \times (1 - FPR(\delta_i))}$$
$$\text{where } \delta = \{\delta_1, \delta_2, \ldots \delta_n\}, \text{ index } i \epsilon \{1, 2, \ldots n\}, 0 \leq \delta_i \leq 1 \quad (4)$$

This can be implemented as a parameter-less (lambda) layer in most of the deep learning frameworks and can be added at the end of CAE. We apply this procedure to the frameworks that are tightly built around results of classification i.e., prediction confidence such as LIME and SHAP. For the other two approaches e-LRP and Counterfactuals, we use the reconstruction based AutoEncoder model without this modification.

4.2. Datasets

The datasets used in our experiments are briefly introduced in this section. Daytime Driver Distraction Dataset (DDDS) [16] is an autonomous driving dataset consisting of image samples that denote the behaviour of 25 drivers by capturing their upper body movements. It consists of 3 classes of anomalies - *Talking, Texting and Operating GPS* with a class containing normal driving patterns of the drivers. The normal samples of 16 drivers at random were used for training and the rest of the normal data along with the anomalous sets were used for testing purposes. MVTec AD dataset [17] is a widely used industrial benchmark dataset with 15 categories *Carpet, Grid, Leather, Tile, Wood, Bottle, Cable, Capsule, Hazelnut,*

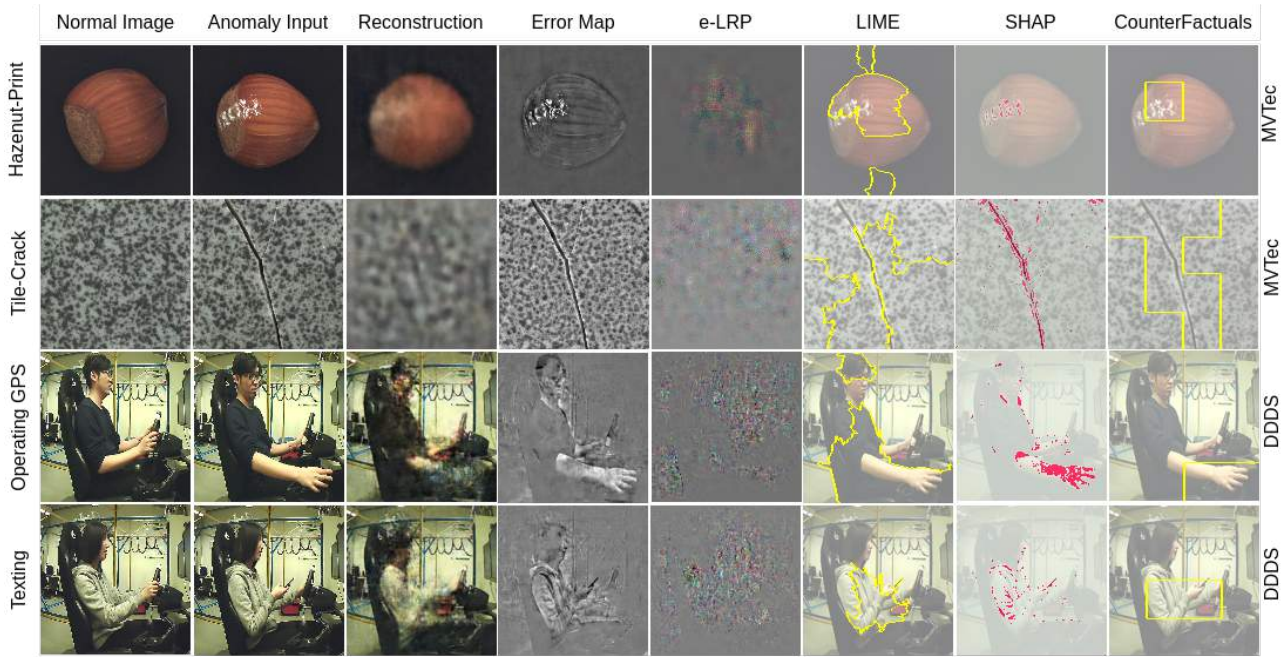


Fig. 1. Comparison of the XAI frameworks on MVTect and DDDS for the anomalies *Print in Hazelnut*, *Crack in a tile*, *Operating GPS* and *Texting* respectively

Metal Nut, Pill, Screw, Toothbrush, Transistor, Zipper of materials collected under the context of industrial inspection and quality control in manufacturing with normal images and with over 70 different defects.

4.3. Discussion of Results

The results of our experiments involving the reconstruction-based anomaly detection performance of convolutional AutoEncoders on MVTec and DDDS datasets are shown in figure 1 with the corresponding normal images and anomalous inputs for comparison. Since there is no established metric to quantify the correctness or accuracy of explanations, we resort to manual visual inspection [2]. It is conspicuous from the results that the explanations from SHAP and Counterfactual are able to precisely indicate the hand movements and the phone responsible for anomalous behaviors like *texting* and *operating GPS* respectively in the DDDS dataset. LIME on the other hand, marks a regional contorted neighborhood without precisely localizing to the anomalous regions. Similar inspection on MVTec shows that SHAP is able to exactly mark the *crack in the tile* and the *print on the hazelnut* thereby precisely indicating the anomalies. Counterfactual method encloses the anomalous regions with a bounding box with a high degree of precision. LIME, although is able to indicate the region, it is not able to localize the the regions as well as SHAP and Counterfactuals. The sub-par performance of LIME can be explained on the basis of the surrogate model used and since it is an approximation of the CAE prediction model, the performance directly depends on the degree of approximation.

The explanation from e-LRP are the pixelated contributions towards the final decision of the framework and are incomprehensible. Since the framework is modified to be a binary classifier with one node (acting as a Sigmoid), explanations from e-LRP are rendered to be ineffective. The better performance of SHAP can be explained on the basis of its working mechanism as it is a more 'flexible' version of counterfactual with pixel-wise or feature-wise pertubations in the input images using random variables instead of region-wise perturbation as in Counterfactuals to determine the individual contributions from the final prediction using the deviation from the original prediction. This allows SHAP to create masks of arbitrary shapes as opposed to rectangular shapes in Counterfactuals although the shape and size of blocks in Counterfactuals can be customized in accordance to the application. Also, comparing the explanations from the modified XAI frameworks with the conventionally used residual reconstruction error maps shows that the latter proves to be ineffective compared to the former owing to the inability to localize to and indicate the anomalous regions precisely. Hence, as hypothesized, residual error maps are not good indicators of model's decisions and XAI frameworks with minor modifications can prove to be better alternatives to understand the decisions of CAEs and also to interpret what is considered by CAEs as 'normal' for visual anomaly detection tasks.

[2]The GitHub repo contains comprehensive results of more data samples tested and detailed descriptions of the datasets

5. CONCLUSION

In this paper, we compared some prevalent explainability frameworks for their potential application in Convolutional AutoEncoders for the task of reconstruction based anomaly detection. We also discussed the modifications required in the design to accomplish the use of these explainability frameworks. Through experiments on two datasets, we have shown that the output of the explored explainability frameworks provide improved insights when compared with the conventional reconstruction residual maps. Explainable AutoEncoders can help in understanding the learning of models in semi/unsupervised tasks and to uncover reclusive patterns from data that are potentially unknown to practitioners. Our follow up works includes the analysis of learning patterns in different types of AutoEncoders for image anomaly detection using XAI methods, with the ultimate goal of paving the way to provide better collaboration between humans and AI.

6. REFERENCES

[1] Manassés Ribeiro, André Eugênio Lazzaretti, and Heitor Silvério Lopes, "A study of deep convolutional auto-encoders for Anomaly Detection in videos," *Pattern Recognition Letters*, vol. 105, pp. 13–22, 2018.

[2] Raghavendra Chalapathy and Sanjay Chawla, "Deep learning for Anomaly Detection: A survey," *arXiv preprint arXiv:1901.03407*, 2019.

[3] Finale Doshi-Velez and Been Kim, "Towards A Rigorous Science of Interpretable Machine Learning," *arXiv preprint arXiv:1702.08608*, 2017.

[4] Quanshi Zhang and Song-Chun Zhu, "Visual Interpretability for Deep Learning: A Survey," *Frontiers of Information Technology Electronic Engineering volume*, p. 27–39, 2018.

[5] Erico Tjoa and Cuntai Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward medical XAI," *IEEE Transactions on Neural Networks and Learning Systems*, p. 1–21, 2020.

[6] Pegah Sagheb Haghighi, Olurotimi Seton, and Olfa Nasraoui, "An Explainable Autoencoder for Collaborative Filtering Recommendation," *CoRR*, vol. abs/2001.04344, 2020.

[7] Liat Antwarg, Bracha Shapira, and Lior Rokach, "Explaining anomalies detected by Autoencoders using SHAP," *CoRR*, vol. abs/1903.02407, 2019.

[8] L. Bergamasco, S. Saha, F. Bovolo, and L. Bruzzone, "An explainable Convolutional Autoencoder model for Unsupervised Change Detection," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 1513–1519, 2020.

[9] Sebastian Bach, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus-Robert Müller, and Wojciech Samek, "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation," *PLOS ONE*, vol. 10, no. 7, pp. 1–46, 07 2015.

[10] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin, ""Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.

[11] Scott M Lundberg and Su-In Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*. 2017, vol. 30, Curran Associates, Inc.

[12] Sandra Wachter, Brent Mittelstadt, and Chris Russell, "Counterfactual explanations without opening the black box: Automated decisions and the GDPR," *Harvard Journal of Law Technology*, vol. 31, pp. 841, 2017.

[13] Sergey Ioffe and Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, 07–09 Jul 2015, vol. 37 of *Proceedings of Machine Learning Research*, pp. 448–456, PMLR.

[14] Abien Fred Agarap, "Deep learning using Rectified Linear Units (relu)," *CoRR*, vol. abs/1803.08375, 2018.

[15] Diederik P. Kingma and Jimmy Ba, "Adam: A method for Stochastic Optimization," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun, Eds., 2015.

[16] Chaojie Ou, Qiang Zhao, Fakhri Karray, and Alaa El Khatib, "Design of an end-to-end dual mode driver distraction detection system," in *Image Analysis and Recognition*, Lecture Notes in Computer Science, pp. 199–207. Springer International Publishing, Cham, 2019.

[17] Paul Bergmann, Kilian Batzner, Michael Fauser, David Sattlegger, and Carsten Steger, "The MVTec Anomaly Detection Dataset: A comprehensive real-world dataset for unsupervised anomaly detection," *International Journal of Computer Vision*, vol. 129, no. 4, pp. 1038–1059, Apr 2021.

HETEROGENEOUS VEHICULAR PLATOONING WITH STABLE DECENTRALIZED LINEAR FEEDBACK CONTROL

Amir Zakerimanesh[1*], *Tony Qiu*[2], *and Mahdi Tavakoli*[1]

[1]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada.
[2]Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada.

ABSTRACT

Platooning which is defined as controlling a group of autonomous vehicles (multiple followers and one leader) to have a desired distance between them while following a desired trajectory has caught on recently in the control engineering discipline. Platooning brings along promising advantages, namely, increasing highway capacity and safety, and reducing fuel consumption. In this paper, using linearized longitudinal dynamic models for each vehicle, we investigate the control problem of vehicular platooning to have all vehicles followed the leader under a constant spacing policy. Under decentralized linear feedback controllers and taking account of heterogeneity in the dynamic models and feedback information to the vehicles, a general dynamic representation for the platoon is obtained. Having this and the proposed controller, stability analysis is developed for any information flow topology (IFT) between vehicles and any number of vehicles. As a case study, a platoon with one leader and two followers is investigated through the proposed strategy, and its stability conditions are provided. Numerical simulations are provided in which the stability range of control gains and the effect of different IFTs on the performance of the platoon are discussed.

Index Terms— Autonomous vehicles, Platoon of vehicles, Stability, Heterogeneity, Information flow topology

1 Introduction

Intelligent transportation systems (ITS) leverage a high level of automation to provide an efficient and safe road transportation. Platooning, which corresponds to travel of a convoy of vehicles with an enforced desired spacing between them, can be subsumed under the ITS discipline. The promise of a reduction in vehicles' fuel consumption due to the decreased aerodynamic drag for back-to-back vehicles [1, 2], and an increased highway capacity and safety [3, 4, 5, 6, 7] warrant more research in this technology. Making sure that all platoon vehicles move at the same velocity as the leader vehicle while keeping a desired spacing among themselves underlies the platoon control problem.

Defining a desired inter-vehicle distance is specified by the spacing policy. Constant distance (CD) policy [8, 9] and constant time headway (CTH) policy [10] are the predominant policies studied in the literature. The CD policy, as its name implies, aims at maintaining a constant distance between consecutive vehicles. In the CTH policy, the spacing between vehicles is dependent on the velocity of the leader and thus no longer constant. Other policies are nonlinear distance policy [11] and delay-based distance policy [12].

From control perspective, dynamics of platoon is characterized by vehicle longitudinal dynamics, information flow topology (IFT), distributed controllers and the spacing policy of the platoon [13, 14]. See [15] to get a quick insight about these components. A platoon is called heterogeneous if the dynamics of the vehicles are not identical.

As linear feedback controllers (LFCs) are concerned, in [16] a decentralized LFC under identical control gains that benefit from position, velocity and acceleration measurements is proposed for a platoon of vehicles, under which the stability conditions for some certain IFTs are derived. In [17], a decentralized LFC is put forward that only utilizes position and velocity feedback signals, and the stability analysis is only applicable for bidirectional and bidirectional-leader IFTs. In [18], a distributed linear control under equal control gains that uses only position signals is devised for the IFT cases that was not addressed in the [16]. In this paper, we use a decentralized LFC with non-identical gains that position, velocity and acceleration of vehicles are fed back into the controllers. In this work and contrary to [16, 18], we incorporate the control gains and the way vehicles communicate with each other directly into the stability analysis of the overall platoon which, therefore, makes it applicable for any IFT, and can specify the stability ranges for the control gains. The adopted method can consider any IFT in the stability analysis, and is applicable for any number of vehicles.

2 Problem formulation

Figure 1 shows a platoon that has $N+1$ (not necessarily identical) vehicles such that the one designated by 0 is the leader

This research is supported by the Government of Alberta's grant to Centre for Autonomous Systems in Strengthening Future Communities (RCP-19-001-MIF). *Correspondence: amir.zakerimanesh@ualberta.ca

vehicle and the others labeled by $1,...,i,i+1,...,N$ are the followers. The distance between the two consecutive vehicles i and $i+1$ is denoted by D_i^{i+1}, and L_i presents the length of the i^{th} follower vehicle. The x axis shows the position of the vehicles during their movement such that x_0 and x_i are the positions of the leader vehicle and the i^{th} follower, respectively. Generally speaking, longitudinal control of a platoon consists of 1) inner force/acceleration control loop, namely feedback linearization (FL) control that compensates for the nonlinear dynamics of the vehicles, and 2) an outer inter-vehicle distance control loop that is responsible for enforcing a desired spacing between the consecutive vehicles within the platoon according to the spacing policy. The FL control is based on the assumption that the vehicle dynamics and its parameters are fully known which means that a perfect nonlinear dynamics cancellation can be achieved. We assume that the FL part has already canceled the dynamics nonlinearities and therefore we will only focus on the inter-vehicle distance control loop. Consider that for platooning, and as far as the leader vehicle is concerned, we only need its position, velocity and acceleration, and it does not undergo any control process. Given that, let the following formulation characterize the dynamics of the i^{th} follower vehicle [19]:

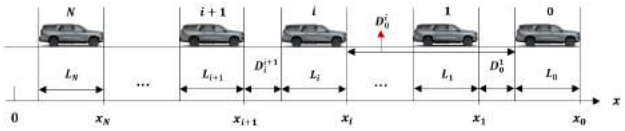


Fig. 1. A platoon with constant inter-vehicle spacing.

$$\dot{a}_i=f_i(v_i,a_i)+g_i(v_i)c_i \qquad i=1,...,N \tag{1}$$

in which v_i and a_i are the velocity and acceleration of the i^{th} follower, and $f_i(v_i,a_i)$ and $g_i(v_i)$ are according to

$$\begin{aligned} f_i(v_i,a_i)&=-\frac{1}{\tau_i}\left(a_i+\frac{\sigma A_i C_{di} v_i^2}{2m_i}+\frac{d_{mi}}{m_i}\right)-\frac{\sigma A_i C_{di} v_i a_i}{m_i} \\ g_i(v_i)&=\frac{1}{\tau_i m_i} \end{aligned} \tag{2}$$

where c_i is the engine input. The parameters $\sigma, A_i, C_{di}, d_{mi}$, m_i, τ_i are specific mass of air, and vehicles' cross sectional area, drag coefficient, mechanical drag, mass, and engine time constant, respectively. Let the engine input c_i be governed by following FL controller:

$$c_i=u_i m_i+0.5\sigma A_i C_{di} v_i^2+d_{mi}+\tau_i \sigma A_i C_{di} v_i a_i \tag{3}$$

substituting which into (1) results in

$$\tau_i \dot{a}_i+a_i=u_i \tag{4}$$

in which u_i is an auxiliary input signal to be designed. Now, let $\mathbf{X}_i \triangleq [x_i, \dot{x}_i, \ddot{x}_i]$ denote the states of the i^{th} follower where $\dot{x}_i=v_i$ and $\ddot{x}_i=a_i$. Thus, given (4), the state-space model for the i^{th} follower can be written as

$$\dot{\mathbf{X}}_i=\mathbf{A}_i\mathbf{X}_i+\mathbf{B}_i u_i=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & -\frac{1}{\tau_i}\end{bmatrix}\mathbf{X}_i+\begin{bmatrix}0\\ 0\\ \frac{1}{\tau_i}\end{bmatrix}u_i \tag{5}$$

where both the vehicles' feedback-linearized dynamics (characterized by $\mathbf{A}_i, \mathbf{B}_i$ and τ_i) and the platoon's controllers (characterized by u_i) are nonidentical, meaning that they are not the same for all the follower vehicles, constituting a heterogeneous platoon. Therefore, the problem formulation and stability analysis will be developed with taking account of heterogeneity in the dynamic models and feedback information to the vehicles.

The objective of designing the controller u_i is to guarantee that when the leader has a constant steady velocity ($\triangleq v_0^s$), the followers' velocities track that leading velocity while desired constant distances ($\triangleq d_i^{i+1}$) are maintained between any two back-to-back vehicles within the platoon. In other words, for $\kappa=1,...,N-1$, the aim is to have

$$\begin{aligned} &v_i(t)=v_0^s(t) \\ &x_\kappa-x_{\kappa+1}=L_\kappa+d_\kappa^{\kappa+1} \quad \equiv \quad D_\kappa^{\kappa+1}=d_\kappa^{\kappa+1} \end{aligned} \tag{6}$$

and to ensure which, we design a distributed controller with non-identical gains as

$$\begin{aligned} u_i&=-\sum_{j\in\mathbb{I}_i}[k_i(x_i-x_j-d_{ij})+b_i(\dot{x}_i-\dot{x}_j)+h_i(\ddot{x}_i-\ddot{x}_j)] \\ d_{ij}&\triangleq -sgn(i-j)\sum_{\kappa=\min(i,j)}^{\max(i,j)-1}\left[l_\kappa+d_\kappa^{\kappa+1}\right] \end{aligned} \tag{7}$$

where $\mathbb{I}_i \subset \{\{0,1,...,N\}-\{i\}\}$ indicates the vehicles from which the vehicle i receives information. Please note that we develop the platooning formulation regardless of the type of communications between the vehicles such that all IFTs can suit properly in our problem development. Having d_i^{i+1} as the desired spacing between the consecutive vehicles and x_0 as the position of the leader vehicle, the desired position and velocity of the i^{th} follower can be defined accordingly as

$$x_i^* \triangleq x_0-\sum_{\kappa=0}^{i-1}\left[l_\kappa+d_\kappa^{\kappa+1}\right], \qquad \dot{x}_i^*=v_0^s=\dot{x}_0^s \tag{8}$$

For conciseness in presentation and ease in later analysis, the state error of the i^{th} follower is defined as $\tilde{x}_i=x_i-x_i^*$ utilizing which readily results in $x_i-x_j=\tilde{x}_i-\tilde{x}_j+d_{ij}$, and subsequently substituting which into the controller (7) gives

$$u_i=-\sum_{j\in\mathbb{I}_i}\left[k_i(\tilde{x}_i-\tilde{x}_j)+b_i(\dot{\tilde{x}}_i-\dot{\tilde{x}}_j)+h_i(\ddot{\tilde{x}}_i-\ddot{\tilde{x}}_j)\right] \tag{9}$$

and plugging (9) in (4) yields

$$\begin{aligned} \dddot{\tilde{x}}_i=&-\frac{|\mathbb{I}_i|k_i}{\tau_i}\tilde{x}_i-\frac{|\mathbb{I}_i|b_i}{\tau_i}\dot{\tilde{x}}_i-\frac{1+|\mathbb{I}_i|h_i}{\tau_i}\ddot{\tilde{x}}_i \\ &+\frac{k_i}{\tau_i}\sum_{j\in\mathbb{I}_i}\tilde{x}_j+\frac{b_i}{\tau_i}\sum_{j\in\mathbb{I}_i}\dot{\tilde{x}}_j+\frac{h_i}{\tau_i}\sum_{j\in\mathbb{I}_i}\ddot{\tilde{x}}_j \end{aligned} \tag{10}$$

which obtained using the facts that $\ddot{x}_i=\ddot{\tilde{x}}_i$ and $\dddot{x}_i=\dddot{\tilde{x}}_i$. Note that $|\mathbb{I}_i|$ is the cardinality of the set $\mathbb{I}_i$. Considering (10), knowing $\tilde{x}_0=\dot{\tilde{x}}_0=\ddot{\tilde{x}}_0=0$, and defining the i^{th} vehicle control gains as $\mathbf{K}_i=[k_i,b_i,h_i]^T$ and platoon state error as $\tilde{\mathbf{X}}_N \triangleq \left[\tilde{x}_1,\dot{\tilde{x}}_1,\ddot{\tilde{x}}_1,...,\tilde{x}_N,\dot{\tilde{x}}_N,\ddot{\tilde{x}}_N\right]^T$, the platoon closed-loop

state-space dynamics model can be characterized by

$$\dot{\tilde{\mathbf{X}}}_N=\tilde{\mathbf{A}}_N\tilde{\mathbf{X}}_N=\begin{bmatrix}\mathbf{A}^*_{11} & \mathbf{A}^*_{12} & \dots & \mathbf{A}^*_{1N}\\ \mathbf{A}^*_{21} & \mathbf{A}^*_{22} & \dots & \mathbf{A}^*_{2N}\\ \vdots & \dots & \ddots & \vdots\\ \mathbf{A}^*_{N1} & \mathbf{A}^*_{N2} & \dots & \mathbf{A}^*_{NN}\end{bmatrix}\tilde{\mathbf{X}}_N \quad (11)$$

where $\tilde{\mathbf{A}}_N$ is overall closed-loop system matrix such that for a given follower i, we have $\mathbf{A}^*_{ii}\triangleq\mathbf{A}_i-|\mathbb{I}_i|\mathbf{B}_i\mathbf{K}_i^T$ and $\mathbf{A}^*_{ij}\triangleq\mathbf{B}_i\mathbf{K}_i^T$. Using $\tilde{\mathbf{A}}_N$, the determinant of the block matrix $s\mathbf{I}_N-\tilde{\mathbf{A}}_N$, which can be obtained analytically [20], will provide the characteristic polynomial of the platoon, using which the stability conditions with respect to the control gains can be obtained. Note that $\mathbf{I}_N$ is the identity matrix of size N, and the closed-loop system would be stable if all the eigenvalues of $\tilde{\mathbf{A}}_N$ are negative. In the rest of paper, we will consider stability conditions for an two-followers platoon.

Case study: stability analysis for $N=2$.
Considering $N=2$, (11) can be written as

$$\dot{\tilde{\mathbf{X}}}_2=\tilde{\mathbf{A}}_2\tilde{\mathbf{X}}_2=\begin{bmatrix}\mathbf{A}_1-|\mathbb{I}_1|\mathbf{B}_1\mathbf{K}_1^T & \mathbf{B}_1\mathbf{K}_1^T\\ \mathbf{B}_2\mathbf{K}_2^T & \mathbf{A}_2-|\mathbb{I}_2|\mathbf{B}_2\mathbf{K}_2^T\end{bmatrix}\tilde{\mathbf{X}}_2 \quad (12)$$

where the platoon would be asymptotically stable if and only if all the eigenvalues of the matrix $\tilde{\mathbf{A}}_2$ are negative. In this respect, the characteristic polynomial of matrix $\tilde{\mathbf{A}}_2$ can be derived by the following determinant:

$$\begin{aligned}&\left|\begin{bmatrix}s\mathbf{I}_3-\mathbf{A}^*_{11} & -\mathbf{A}^*_{12}\\ -\mathbf{A}^*_{21} & s\mathbf{I}_3-\mathbf{A}^*_{22}\end{bmatrix}\right|\\ &=|s\mathbf{I}_3-\mathbf{A}^*_{11}|\left|(s\mathbf{I}_3-\mathbf{A}^*_{22})-\mathbf{A}^*_{21}(s\mathbf{I}_3-\mathbf{A}^*_{11})^{-1}\mathbf{A}^*_{12}\right|\end{aligned} \quad (13)$$

deriving which presents the characteristic polynomial $as^6+bs^5+cs^4+ds^3+es^2+fs^1+g$ in which the coefficients are according to the following formulas.

$$\begin{aligned}&a=\tau_1\tau_2 \qquad b=\tau_1(1+h_2|\mathbb{I}_2|)+\tau_2(1+h_1|\mathbb{I}_1|)\\ &c=\tau_1b_2|\mathbb{I}_2|+(1+h_1|\mathbb{I}_1|)(1+h_2|\mathbb{I}_2|)+\tau_2b_1|\mathbb{I}_1|-h_1h_2\\ &d=\tau_1k_2|\mathbb{I}_2|+b_2|\mathbb{I}_2|(1+h_1|\mathbb{I}_1|)+b_1|\mathbb{I}_1|(1+h_2|\mathbb{I}_2|)\\ &\quad+\tau_2k_1|\mathbb{I}_1|-b_1h_2-b_2h_1\\ &e=k_2|\mathbb{I}_2|(1+h_1|\mathbb{I}_1|)+b_1|\mathbb{I}_1|b_2|\mathbb{I}_2|+k_1|\mathbb{I}_1|(1+h_2|\mathbb{I}_2|)\\ &\quad-k_2h_1-b_1b_2-h_2k_1\\ &f=b_1|\mathbb{I}_1|k_2|\mathbb{I}_2|+k_1|\mathbb{I}_1|b_2|\mathbb{I}|_2-k_2b_1-b_2k_1\\ &g=k_1|\mathbb{I}_1|k_2|\mathbb{I}_2|-k_1k_2\end{aligned} \quad (14)$$

and if the first follower does not receive information from the second follower, or vice versa, then we will have $\mathbf{A}^*_{12}=0$ or $\mathbf{A}^*_{21}=0$, respectively. Thus, the coefficients would be

$$\begin{aligned}&a=\tau_1\tau_2 \qquad b=\tau_1(1+h_2|\mathbb{I}_2|)+\tau_2(1+h_1|\mathbb{I}_1|)\\ &c=\tau_1b_2|\mathbb{I}_2|+\tau_2b_1|\mathbb{I}_1|+\tau_1(1+h_2|\mathbb{I}_2|)+\tau_2(1+h_1|\mathbb{I}_1|)\\ &d=\tau_1k_2|\mathbb{I}_2|+b_2|\mathbb{I}_2|(1+h_1|\mathbb{I}_1|)+b_1|\mathbb{I}_1|(1+h_2|\mathbb{I}_2|)+\tau_2k_1|\mathbb{I}_1|\\ &e=k_2|\mathbb{I}_2|(1+h_1|\mathbb{I}_1|)+b_1|\mathbb{I}_1|b_2|\mathbb{I}_2|+k_1|\mathbb{I}_1|(1+h_2|\mathbb{I}_2|)\\ &f=b_1|\mathbb{I}_1|k_2|\mathbb{I}_2|+k_1|\mathbb{I}_1|b_2|\mathbb{I}|_2 \qquad g=k_1|\mathbb{I}_1|k_2|\mathbb{I}_2|\end{aligned} \quad (15)$$

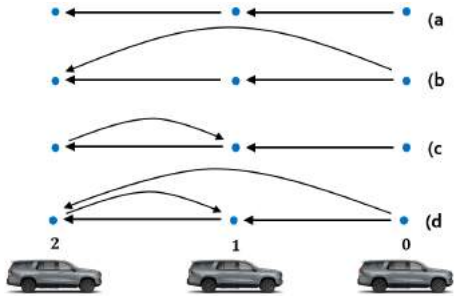


Fig. 2. Schematic of different IFTs between the vehicles in the one-leader-two-followers platoon.

Now, having (14)-(15) and using Routh–Hurwitz criterion, the stability conditions can be obtained as follows.

$$\begin{aligned}&1.\ a,b,c,d,e,f,g>0 \quad 2.\ ad-bc\le 0 \quad 3.\ d(ad-bc)\le b(af-be)\\ &4.\ (ad-bc)\left[b^2g+f(ad-bc)\right]\le(af-be)[d(ad-bc)-b(af-be)]\\ &5.\ \left(b^2g+f(ad-bc)\right)\left[(ad-bc)\left[b^2g+f(ad-bc)\right]\right.\\ &\quad\left.-(af-be)[d(ad-bc)-b(af-be)]\right]\\ &\quad\ge bg[d(ad-bc)-b(af-be)]^2\end{aligned} \quad (16)$$

3 Simulation Results

In this section, simulation results are provided to evaluate the stability conditions for different IFTs that are depicted in Fig. (2). For simulations, we consider a velocity trajectory for the leader vehicle (see Fig. 4) and choose the vehicles' initial velocities and accelerations equal to zero. Also, the vehicles' length are the same and equal to 4 m, and vehicles' initial positions are selected as $x_0(0)=0$ m, $x_1(0)=-10$ m, and $x_2(0)=-20$ m. As you can see in the Fig. 4, the v_0^s velocities for the leader vehicle are 30 m/s (its maximum value) persisting for 12 s, and 0 m/s that is associated with the time the leader vehicle brakes and stands still. Furthermore, we choose $d_i^{i+1}=10$ m as the desired spacing between the vehicles.

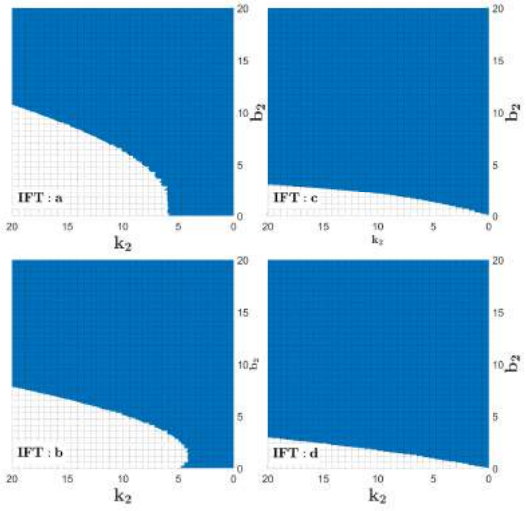


Fig. 3. The stability area (the blue area) with respect to the control gains k_2 and b_2 for different IFTs sketched in Fig. 2.

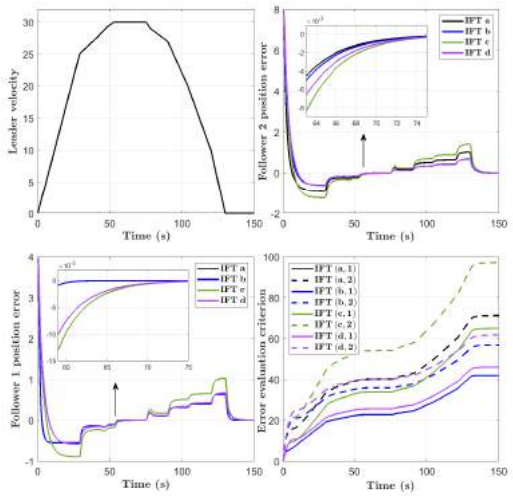


Fig. 4. Error signals of the followers for the different IFTs.

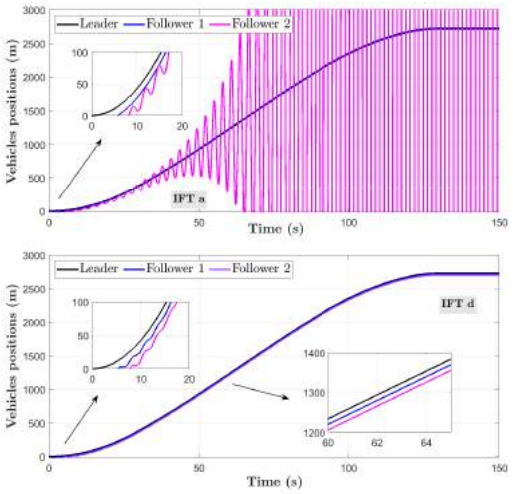


Fig. 5. Vehicles' positions using control gains k_1=3, b_1=5, h_1=1, k_2=10, b_2=2, and h_2=1, and IFTs a and d.

First, we assume that τ_1=τ_2=0.5 s, and the controller gains of all the vehicles are the same, i.e., k_1=k_2, b_1=b_2 and h_1=h_2=1. Based upon the stability conditions given in the work [16] and for IFT c illustrated in Fig. 2, we assign k_1=k_2=3, b_1=b_2=5, and h_1=h_2=1. Having k_1,b_1,h_1, we choose h_2=h_1 and let k_2 and b_2 to be selected within the stability conditions given in (16). Regarding (14)-(15), this time we will find stability conditions with respect to the control gains k_2 and b_2 and for the four IFTs in Fig. 2. The results for the different IFTs are depicted in Fig. 3. The stability areas are shown in Fig. 3. As you can see, by comparing the stability areas of IFTs a and b, or IFTs c and d, or IFTs a and c, and or IFTs b and d, an additional communication channel between the vehicles makes the stability area larger. The IFT a has the smallest stability area and the IFT d has the largest.

In order to draw an analogy between the controller performances in different IFTs, using root locus analysis for a given plausible k_2 or b_2 that belongs to all the stability areas of Fig. 3, we assign k_2=2.5 and b_2=10. Therefore, the control gains become k_1=3, k_2=2.5, b_1=5, b_2=10, and h_1=h_2=1. Note here τ_1=0.5 and τ_2=0.5 are chosen for engines time constants. So, using this controller, the results for the different IFTs are shown in Fig. 4 in which, for instance, IFT $(a,2)$ indicates the position error for the second follower and implies that the controller is utilized within the IFT of case a represented in Fig. 2. Note that the position error for the i^{th} follower is defined as $e_i(t)$=$x_i(t)$−$x_i^*(t)$. Investigating the simulation results for the different IFTs, we can see that when the leader has a constant steady velocity, the followers' position errors asymptotically converge to zero. Also, in IFTs b and d, in which both the first and second followers receive information from the leader, the error signal exhibits better damping behavior that can come in handy when, for instance, we want to enforce small desired spacing between the vehicles. To shed more light on the damping behavior, let the following formula be defined as the *error evaluation criterion* (EEC) for the transient behavior of the error signals of the followers.

$$EEC_i \triangleq \int_0^t |e_i(t)|dt \tag{17}$$

regarding which the results for the followers within the given IFTs are shown in Fig. 4. It is possible to see that the IFTs b and d provide better performance for the platoon respecting EEC measure. Moreover, making a comparison between the IFTs b and d, we can see that the communication from the second follower to the first follower has increased the settling time and so the convergence occurs slower.

Fig. 5 shows the positions of vehicles for the given velocity of the leader and for the two IFTs a and d. As obvious from Fig. 3, for k_2=10 and b_2=2, the platoon of the IFT a would be unstable and the platoon of the IFT d would be stable. Accordingly, in Fig. 5, using the IFT d, the desired distances between the vehicles are maintained, however, in the IFT a the system is unstable and numerous collisions occur.

4 Conclusion

In this paper, using a decentralized linear feedback controller with non-identical gains, a state-space model for the heterogeneous platoon was obtained. We developed the problem in such a way as to could incorporate any IFT into the stability analysis. Thus, for any number of vehicles, using the characteristic polynomial of the closed-loop system, the Routh–Hurwitz criterion will present the stability conditions of the platoon. As a case study, the simulation results were provided for an two-followers platoon, and the effect of the different IFTs on the system performance were discussed. It was shown that, more communication between the vehicles can provide more flexibility in the selection of control gains that satisfy the stability conditions. The results also showed that using feedback signals of the leader in the both followers' controllers can offer better performance for the platoon.

5 References

[1] Assad Alam, Bart Besselink, Valerio Turri, Jonas Mårtensson, and Karl H Johansson, "Heavy-duty vehicle platooning for sustainable freight transportation: A cooperative method to enhance safety and efficiency," *IEEE Control Systems Magazine*, vol. 35, no. 6, pp. 34–56, 2015.

[2] Christophe Bonnet and Hans Fritz, "Fuel consumption reduction in a platoon: Experimental results with two electronically coupled trucks at close spacing," Tech. Rep., SAE Technical Paper, 2000.

[3] Mudasser Seraj, Jiangchen Li, and Zhijun Qiu, "Modeling microscopic car-following strategy of mixed traffic to identify optimal platoon configurations for multiobjective decision-making," *Journal of Advanced Transportation*, vol. 2018, 2018.

[4] Jiwen Tan, Tangtao Yang, Yi Zhang, and Tony Z Qiu, "Evaluation of vehicles' platooning on expressways based on v2x," in *2019 5th International Conference on Transportation Information and Safety (ICTIS)*. IEEE, 2019, pp. 369–375.

[5] Jiangchen Li, Chen Qiu, Mudasser Seraj, Liqun Peng, and Tony Z Qiu, "Platoon priority visualization modeling and optimization for signal coordination in the connected vehicle environment," *Transportation research record*, vol. 2673, no. 5, pp. 36–48, 2019.

[6] Steven E Shladover, Dongyan Su, and Xiao-Yun Lu, "Impacts of cooperative adaptive cruise control on freeway traffic flow," *Transportation Research Record*, vol. 2324, no. 1, pp. 63–70, 2012.

[7] Fei-Yue Wang, "Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 3, pp. 630–638, 2010.

[8] Petros A Ioannou and Cheng-Chih Chien, "Autonomous intelligent cruise control," *IEEE Transactions on Vehicular technology*, vol. 42, no. 4, pp. 657–672, 1993.

[9] DVAHG Swaroop and J Karl Hedrick, "Constant spacing strategies for platooning in automated highway systems," 1999.

[10] Jing Zhou and Huei Peng, "Range policy of adaptive cruise control vehicles for improved flow stability and string stability," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 229–237, 2005.

[11] Gábor Orosz, "Connected cruise control: modelling, delay effects, and nonlinear behaviour," *Vehicle System Dynamics*, vol. 54, no. 8, pp. 1147–1176, 2016.

[12] Bart Besselink and Karl H Johansson, "String stability and a delay-based spacing policy for vehicle platoons subject to disturbances," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4376–4391, 2017.

[13] Srdjan S Stankovic, Milorad J Stanojevic, and Dragoslav D Siljak, "Decentralized overlapping control of a platoon of vehicles," *IEEE Transactions on Control Systems Technology*, vol. 8, no. 5, pp. 816–832, 2000.

[14] Peter Seiler, Aniruddha Pant, and Karl Hedrick, "Disturbance propagation in vehicle strings," *IEEE Transactions on automatic control*, vol. 49, no. 10, pp. 1835–1842, 2004.

[15] Shuo Feng, Yi Zhang, Shengbo Eben Li, Zhong Cao, Henry X Liu, and Li Li, "String stability for vehicular platoon control: Definitions and analysis methods," *Annual Reviews in Control*, vol. 47, pp. 81–97, 2019.

[16] Yang Zheng, Shengbo Eben Li, Jianqiang Wang, Dongpu Cao, and Keqiang Li, "Stability and scalability of homogeneous vehicular platoon: Study on the influence of information flow topologies," *IEEE Transactions on intelligent transportation systems*, vol. 17, no. 1, pp. 14–26, 2015.

[17] Ali Ghasemi, Reza Kazemi, and Shahram Azadi, "Stable decentralized control of a platoon of vehicles with heterogeneous information feedback," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4299–4308, 2013.

[18] Shengbo Eben Li, Xiaohui Qin, Yang Zheng, Jianqiang Wang, Keqiang Li, and Hongwei Zhang, "Distributed platoon control under topologies with complex eigenvalues: Stability analysis and controller synthesis," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 1, pp. 206–220, 2017.

[19] S Huang and Weili Ren, "Longitudinal control with time delay in platooning," *IEE Proceedings-Control Theory and Applications*, vol. 145, no. 2, pp. 211–217, 1998.

[20] Philip D Powell, "Calculating determinants of block matrices," *arXiv preprint arXiv:1112.4379*, 2011.

TRUSTWORTHY ADAPTATION WITH FEW-SHOT LEARNING FOR HAND GESTURE RECOGNITION

Elahe Rahimian[†]*, Soheil Zabihi*[‡]*, Amir Asif*[‡]*, S. Farokh Atashzar*[††]*, and Arash Mohammadi*[†]

[†]Concordia Institute for Information System Engineering, Concordia University, Montreal, QC, Canada
[‡]Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada
[††]Electrical & Computer Engineering, Mechanical & Aerospace Engineering, New York University, USA

ABSTRACT

This work is motivated by potentials of Deep Neural Networks (DNNs)-based solutions in improving myoelectric control for trustworthy Human-Machine Interfacing (HMI). In this context, we propose the Trustworthy Few Shot-Hand Gesture Recognition (TFS-HGR) framework as a novel DNN-based architecture for performing Hand Gesture Recognition (HGR) via multi-channel surface Electromyography (sEMG) signals. The main objective of the TFS-HGR framework is to employ Few-Shot Learning (FSL) formulation with a focus on transferring information and knowledge between source and target domains (despite their inherit differences) to address limited availability of training data. The NinaPro DB5 dataset is used for evaluation purposes. The proposed TFS-HGR achieves a performance of 83.17% for new repetitions with few-shot observations, i.e., 5-way 10-shot classification. Moreover, the TFS-HGR with the accuracy of 75.29% also generalize to new gestures with few-shot observations, i.e., 5-way 10-shot classification.

***Index Terms*—** Few-Shot Learning (FSL), surface Electromyographic (sEMG), Hand Gesture Recognition (HGR), Attention Mechanism, Temporal Convolution

1. INTRODUCTION

Recent developments in rehabilitation and assistive technologies combined with advances in Machine Learning (ML) and Deep Neural Networks (DNNs) have led to the development of improved myoelectric prosthesis control systems. The surface electromyographic (sEMG) signals are especially used to classify limb movements and improve the quality of myoelectric prosthesis control systems, which enhance living conditions of hand amputated individuals. The sEMG signals are collected by wrapping sensors around the skin to measure the muscle fibers' action potentials. Because various applications of hand gesture recognition, such as prosthetic control, involve continuous and long-term use, it is essential to develop a learning algorithm that incorporates prior knowledge and experience gathered from the source domain with new information. In addition, there is a need to develop adaptive learning algorithms with the aim of designing a classifier, which can be used for new gestures based on only a few observations through the fast learning approach. This is a challenging task since many factors such as electrode shifts and muscle fiber lengthening/shortening can affect the gathered sEMG signals. Moreover, the neurophysiology differences between users result in more inconsistencies between different conditions [1, 2]. In this context, by taking advantage of Few-Shot Learning (FSL) the gathered experience from several source can speed up the recognition process for a new gesture. Thus, the learning process does not start from the beginning every time for an unseen gesture. In this regard, the paper focuses on hand gesture classification [3–13] using FSL to reduce the critical challenge of variability in the characteristics of sEMG signals. In other words, we propose a framework known as Trustworthy Few Shot-Hand Gesture Recognition (TFS-HGR) that utilizes a combination of temporal convolutions and attention mechanisms. The proposed TFS-HGR framework allows a myoelectric controller that has built based on background data to adapt the variations in stochastic characteristics of sEMG signals. The adaptation can be achieved with a small number of new observations making it suitable for clinical implementations and practical applications.

The paper is organized as follows: Section 1.1 provides an overview of relevant literature. In Section 2, we describe the proposed TFS-HGR architecture. Experimental results and different evaluation scenarios with dataset used in development of the proposed architecture are presented in Section 3. Finally, Section 4 concludes the paper.

1.1. Literature Review

A common strategy used to identify hand movements with DNN-based models is to consider 2/3 of the gesture trials of each user for the training set, while the rest of trials constitute the test set [5–13]. Despite extensive advances in myoelectric prostheses in recent DNN-based models, their performance is significantly reduced when data are limited. Collecting large data sets for training may be possible in research laboratories, but this is not a practical method for real applications. Therefore, the use of domain adaptation algorithms is a fundamen-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

tal step in bridging the gap between real-world practice and academic achievement in the laboratory. In this regard, some researches such as [14, 15] utilize Transfer Learning (TL)-based algorithms to take advantage the knowledge obtained from source users and accelerate the learning process for new target users. In addition, Reference [16] adopts a TL-based algorithm and provides an adaptive calibration of electrode array shifts which enhances the robustness of the myoelectric control system. Moreover, in References [17, 18], the authors use a long-term recording of sEMG data to train and test various DNN-based architectures to provide a robust algorithm against non-stationary nature of sEMG signals varying over days.

Most of the existing techniques (such as [17, 18]) have been tested on data of limited dimension, with respect to the number of users and/or the number of hand gestures. Therefore, their generalizability to a larger population with an increased number of gestures has not been evaluated. Therefore, It is not clear how these models would perform on scenarios with a larger number of users and gestures. Therefore, despite the recent targeted focus on improving robustness in DNN-based myoelectric control models, more research is still needed, particularly on the design of new classification models, a unified framework for FSL, and domain adaptation algorithms. The paper, inspired by [19], makes the following significant contributions to the growing research in this area:

- by proposing a FSL framework that allows the myoelectric controller that has been built on background data from source gestures/repetitions, to adapt to the changes in the EMG signal characteristics from a target gesture/repetition with minimal efforts and a few observations, making it suitable for clinical settings and practical applications.
- By proposing the TFS-HGR framework, which utilizes a combination of temporal convolutions and attention mechanisms, we provide a novel venue for adopting few-shot learning, to not only reduce the training time, but also to eventually mitigate the significant challenge of variability in the characteristics of sEMG signals.

2. THE PROPOSED TFS-HGR ARCHITECTURE

Fast learning is an indication of human intelligence, which involves recognizing an object just by looking at a few examples or quickly repeating an action several times. There is a need to develop adaptive learning methods focusing on designing classifications for myoelectric control systems, which can be employed for new gestures based on only a few observations through a fast learning approach. In this regard, by inspiring from Reference [19], we propose the FSL-based architecture which utilized the temporal convolutions along with attention mechanism. In the following, first, we briefly describe the FSL formulation, and then provide more details about the building blocks of TFS-HGR architecture.

2.1. The FSL Formulation

Within FSL formulation, we are dealing with tasks where each task $\mathcal{T}_j$ corresponds to a dataset $\mathcal{D}_j$ that splits into two parts, a support-set ${\mathcal{D}_j}^{support}$, and a query-set ${\mathcal{D}_j}^{query}$. The ${\mathcal{D}_j}^{support}$ consists of a sequence of observation-label pairs, while the ${\mathcal{D}_j}^{query}$ consists of an unlabelled observation. Therefore, based on the labels in the ${\mathcal{D}_j}^{support}$, the network is going to predict the missing label in the ${\mathcal{D}_j}^{query}$. The parameters of the network are optimized using the loss between the prediction and the ground truth label in the ${\mathcal{D}_j}^{support}$ over mini-batches of tasks. During the test procedure, unseen tasks are sampled from a different task distribution (i.e., $\widetilde{\mathcal{T}}_j \sim p\left(\widetilde{\mathcal{T}}\right)$) that is similar in nature to $p\left(\mathcal{T}\right)$.

More specifically, we follow reference [19] and define the N-way k-shot classification task as such that the support-set consists of k examples for each of N unique classes. The N classes are randomly selected from the overall dataset. Each observation in the support-set is concatenated with its corresponding label constituting the $N \times k$ time-step for the network. These $N \times k$ observation-label pairs are followed by an unlabelled observation which is sampled from one of the N classes. This observation is concatenated with a null label. The network is supposed to predict the missing label of the last $(N \times k + 1)^{th}$ time-step.

2.2. The Building Blocks of TFS-HGR Architecture

The proposed TFS-HGR architecture consists of several structural blocks whose algorithms are provided in Algorithm 1, Algorithm 2, and Algorithm 3. Each block receives an input with size $(C_{in} \times l)$, where C_{in} represents the input dimensionality and l is the length of sequence which is equal to $(N \times k+1)$. It should be noted that before feeding the input to the proposed TFS-HGR framework, it becomes a feature vector. For this target, we used an *Embedding Module*.

More specifically, For performing HGR, it is common to segment sEMG data through a window before feeding it as input to neural networks. Therefore, in this work, before feeding the sEMG data, it is segmented by a window with size $W \times N$, where W shows the length of the window and N indicates the number of sensors. Then, an Embedding Module is used to convert the input with size $W \times N$ to a feature vector. This feature vector is used as input for subsequent modules in the proposed TFS-HGR network. In this paper, we adopted 2 Embedding Module. The first one, is using the Fully Connected (FC) layers for extracting the feature vector. The second case is the use of Long Short Term Memory (LSTM) as the Embedding Module for extracting this feature vector.

The next building block is *The TemporalBlock Module* which consists of two Dilated Causal 1D-Convolutions with dilation factor d, kernel size k_S, and f filters. Each Dilated Causal 1D-Convolution is followed by ReLU activation func-

Table 1: Experiment 1: 5-way 1-shot, 5-way 5-shot, and 5-way 10-shot classification accuracies based on *new repetitions with few-shot observation*. In this experiment, we adopted two different Embedding Modules: (i) FC Embedding and (ii) LSTM Embedding

The Embedding Module	5-way Accuracy (%) ± STD		
	1-shot	5-shot	10-shot
FC Embedding	66.75 ± 0.14	79.48 ± 0.13	79.56 ± 0.15
LSTM Embedding	69.61 ± 0.15	78.75 ± 0.14	83.17 ± 0.15

Algorithm 1 THE TEMPORALBLOCK MODULE

function: TemporalBlock(input, dilation factor d, kernel size k_S, number of filters f):
output1 = CausalConv(input, d, k_S, f)
activation1 = relu(output1)
output2 = CausalConv(activation1, d, k_S, f)
activation2 = relu(output2)
return concat(input, activation2)

Algorithm 2 THE TEMPORALCONVNET MODULE

function: TemporalConvNet(input, sequence length $l = (N \times k + 1)$, kernel size k_S, number of filters f):
$Z = \lceil \log_2 l \rceil$
for i in $0, .., Z-1$ **do**
    input = TemporalBlock(input, 2^i, k_S, f)
end for
return input

tion. Finally, the input and output are concatenated (Algorithm 1).

The TemporalConvNet Module consists of $Z = \lceil \log_2 l \rceil$ TemporalBlock Modules whose dilation factor increases exponentially (Algorithm 2).

The Attention Module performs similarity measurement between keys and queries which more details are provided in Algorithm 3 and Reference [20].

Algorithm 3 THE ATTENTION MODULE

function: Attention(input, key size d_k, value size d_v):
$\boldsymbol{K}$ = affine(input, d_k)
$\boldsymbol{Q}$ = affine(input, d_k)
$\boldsymbol{V}$ = affine(input, d_v)
logits = matmul($\boldsymbol{Q}$, transpose($\boldsymbol{K}$))
probs = softmax($\frac{\text{logits}}{\sqrt{d_k}}$)
output = matmul(probs, $\boldsymbol{V}$)
return concat(input, output)

3. EXPERIMENTS AND RESULTS

In this section, the database used to evaluate the proposed architecture is first described. Then, the various experiments will be discussed in the following sections. It is noteworthy to mention that we evaluated the proposed model for two scenarios. In all the experiments, for training purpose, the Adam optimizer with learning rate of 0.0001 was used. The loss function is obtained using Cross-entropy between the predicted and the truth label of final example in each $\mathcal{T}_j$. The mini-batch with size 64 also was used in the all experiments.

3.1. Database

In this paper, the 5th Ninapro database [21] referred to as the DB5 is used. DB5 dataset is recorded with two Thalmic Myo armbands including 16 active singledifferential wireless electrodes, recording muscular activity at a rate of 200 Hz. More specifically, the DB5 dataset consists of signals collected from 10 intact-limb users performing 52 movements including basic movements of the fingers, isometric, isotonic hand configurations, basic wrist movements, and grasping and functional movements. Each movement in the DB5 dataset is repeated 6 times, each lasting for 5 seconds followed by 3 seconds of rest. The DB5 dataset was presented in three sets of exercises which more details are provided in [21]. It is noteworthy to say that to follow the same criteria in [14, 15] and also have a fair comparison, in this paper, we only consider the lower armband in DB5.

3.2. Experiment 1

In this experiment, we evaluate the performance of the proposed TFS-HGR architecture for *new repetitions* with few-shot examples for DB5 dataset. We considered $2/3$ of the gesture repetitions of each subject for training and validation purpose. The remaining repetitions constitute the test set. Table 1 shows the performance of model when FSL is applied for HGR task. In this experiment, we used two Embedding Modules, i.e., FC and LSTM Embedding. It can be observed that the Embedding Module could affect the overall accuracy. Moreover, by increasing the number of shots, the accuracy will be improved. It is noteworthy to mention that the performance of the proposed TFS-HGR architecture for this experiment is evaluated based on the second set of exercises DB5, which includes 17 movements.

3.3. Experiment 2

In this experiment, it is shown that the proposed TFS-HGR model can be used for new movements based on only a few

Table 2: Experiment 2: 5-way 1-shot, 5-way 5-shot, and 5-way 10-shot classification accuracies based on *new gestures with few-shot observation*. In this experiment, we adopted two different Embedding Modules: (i) FC Embedding and (ii) LSTM Embedding

The Embedding Module	**5-way Accuracy (%) ± STD**		
	1-shot	**5-shot**	**10-shot**
FC Embedding	40.23 ± 0.14	65.95 ± 0.15	75.29 ± 0.15
LSTM Embedding	39.56 ± 0.15	65.14 ± 0.14	74.91 ± 0.16

examples. More precisely, the test set consists of completely new gestures. This is a challenging task because unlike the Experiment 1, here, the probability distribution of the test set is not the same as training set. In other words, the model leverage the experience from source gestures to adapt to the variations in the new gestures in the target. Table 2 shows the efficiency of the proposed model when we had out-of-sample movements in the test set. The model predicted unknown class distributions in scenarios where few observations from the target distribution were available. It is worth to mention that in this experiment, 52 movements in DB5 are used. More specifically, 68% of total gestures are considered for the training purpose, while the validation set consists of 12% of the total gestures. The test set consists of the remaining gestures (20% of the total gestures).

4. CONCLUSION

In this work, a novel FSL architecture is proposed for HGR through sEMG signals. The proposed TFS-HGR architecture can generalize after seeing very few examples of each class, and then use the obtained knowledge for predicting the new classes. The ability to use the obtained knowledge during the training and adapt to the new classes is a feature of the proposed architectures. In this paper, we showed that by taking advantage of FSL, the proposed architecture can predict the new gestures in the target domain, and the learning does not have to start from the beginning. This paper is a huge step toward proposing and developing FSL framework for a target domain when there is a distribution mismatch with the source domain, e.g., new users or electrode shift. In this regard, the focus of the future works is developing more robust domain adapting algorithms for HGR tasks.

5. REFERENCES

[1] C. Castellini, *et al.*, "Peripheral Machine Interfaces: Going Beyond Traditional Surface Electromyography," *Frontiers in Neurorobotics*, vol. 8, p. 22, 2014.

[2] D. Farina, *et al.*, "The Extraction of Neural Information from the Surface EMG for the Control of Upper-limb Prostheses: Emerging Avenues and Challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol 22, no.4, pp. 797-809, 2014.

[3] K. Lai and S. Yanushkevich, "An Ensemble of Knowledge Sharing Models for Dynamic Hand Gesture Recognition," *Int. Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1-7.

[4] K. Lai and S. N. Yanushkevich, "CNN+RNN Depth and Skeleton based Dynamic Hand Gesture Recognition," *International Conference on Pattern Recognition (ICPR)*, 2018, pp. 3451-3456.

[5] E. Rahimian, S. Zabihi, S. F. Atashzar, A. Asif, and A. Mohammadi, "Surface EMG-Based Hand Gesture Recognition via Hybrid and Dilated Deep Neural Network Architectures for Neurorobotic Prostheses," *Journal of Medical Robotics Research*, pp. 1-12, 2020.

[6] E. Rahimian, S. Zabihi, F. Atashzar, A. Asif, A. Mohammadi, "XceptionTime: Independent Time-Window XceptionTime Architecture for Hand Gesture Classification," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020.

[7] E. Rahimian, S. Zabihi, S. F. Atashzar, A. Asif, A. Mohammadi, "sEMG-Based Hand Gesture Recognition via Dilated Convolutional Neural Networks," *IEEE Gloabl Confeerence on Signal and Information Processing (GlobalSIP)*, 2019.

[8] E. Rahimian, S. Zabihi, A. Asif, D. Farina, S.F. Atashzar, and A. Mohammadi "FS-HGR: Few-shot Learning for Hand Gesture Recognition via ElectroMyography.," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 2021. In Press.

[9] E. Rahimian, S. Zabihi, F. Atashzar, A. Asif, A. Mohammadi, "Few-Shot Learning for Decoding Surface Electromyography for Hand Gesture Recognition," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2021, pp. 1300-1304.

[10] W. Wei, *et al.*, "Surface Electromyography-based Gesture Recognition by Multi-view Deep Learning," *IEEE Trans. Biomedical Eng.*, vol. 66, no. 10, pp. 2964-2973, 2019.

[11] Y. Hu, *et al.*, "A Novel Attention-based Hybrid CNN-RNN Architecture for sEMG-based Gesture Recognition," *PloS one*, vol. 13, no. 10, p.e0206049, 2018.

[12] L. Chen, J. Fu, Y. Wu, H. Li, and B. Zheng, "Hand Gesture Recognition Using Compact CNN Via Surface Electromyography Signals," *Sensors*, vol. 20, no.3, p. 672, 2020.

[13] W. Wei, *et al.* "A Multi-stream Convolutional Neural Network for sEMG-based Gesture Recognition in Muscle-computer Interface," *Pattern Recognition Letters*, 119, pp. 131-138, 2019.

[14] U. Ct-Allard, *et al.* "Deep Learning for Electromyographic Hand Gesture Signal Classification using Transfer Learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760-771, 2019.

[15] U. Ct-Allard, *et al.*, "Transfer learning for SEMG Hand Gestures Recognition Using Convolutional Neural Networks," *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2017, pp. 16631668.

[16] X. Zhang, L. Wu, B. Yu, X. Chen, and X. Chen, "Adaptive Calibration of Electrode Array Shifts Enables Robust Myoelectric Control," *IEEE Trans. Biomedical Eng.*, 2019.

[17] M. Zia ur Rehman, *et al.*, "Multiday EMG-based Classification of Hand Motions with Deep Learning Techniques," *Sensors*, vol. 18, no. 8, p. 2497, 2018.

[18] Y. Du, W. Jin, W. Wei, Y. Hu, and W. Geng, "Surface Emg-based Inter-session Gesture Recognition Enhanced by Deep Domain Adaptation," *Sensors*, vol. 17, no. 3, p. 458, 2017.

[19] N. Mishra, M. Rohaninejad, X. Chen, and P. Abbeel, "A Simple Neural Attentive Meta-Learner," *arXiv preprint arXiv:1707.03141*, 2017.

[20] A. Vaswani, N. Shazeer, J. Uszkoreit, L. Jones, A. Gomez N., L. Kaiser, and I. Polosukhin, "Attention is All You Need," *arXiv preprint arXiv:1706.03762*, 2017a.

[21] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Muller, and M. Atzori, "Comparison of Six Electromyography Acquisition Setups on Hand Movement Classification Tasks," *PLoS ONE*, vol. 12, no. 10, pp. 1-7, 2017.

THERMAL FACE IMAGE GENERATOR

Xingdong Cao, Kenneth Lai, Svetlana Yanushkevich, Michael Smith

Department of Electrical and Software Engineering, University of Calgary, Alberta, T2N 1N4 Canada

ABSTRACT

This work addresses two image-to-image translation tasks. The first task is to convert a visible face image into a thermal face image (V2T) and the second task is to convert a thermal face image into another thermal face image with a given target temperature (T2T). We propose to use conditional generative adversarial networks to solve the two tasks. We train our models using Carl and SpeakingFaces Datasets, and use SSIM to measure the performance of our models. The SSIM of the generated thermal images reach 0.82 and 0.84 for the V2T and T2T tasks respectively.

Index Terms— Generative adversarial networks, image-to-image translation, thermal image generation

1. INTRODUCTION

The visible and thermal spectra provide two different observations of an object. The most recent researches focus on the visible spectrum where it is easy to capture face biometrics. In contrast, many biometrics details are lost from the thermal spectrum, while it is straight-forward to extract temperature information and capture more scene information in low light than with visible spectrum cameras. The question this paper wants to answer is "*Is it possible to convert a face image taken in visible spectrum into an image taken in the thermal spectrum (V2T)*"? Such an outcome of can be applied into many fields such as healthcare, airport surveillance, human body simulation and medical science.

In addition, we want to know what a face will look like in the thermal spectrum if the face has a different temperature. So, the second problem this paper will solve is "*How to convert a thermal face image into another thermal image with a target temperature (T2T)*"? With this, we can simulate the thermal pattern without adequate thermal dataset of a living body.

The main contribution of this paper is that, we build a deep learning model to solve these two image-to-image translation tasks. We use a conditional generative adversarial network (cGAN) [1] that is conditioned on some input to generate the output images. Our cGAN consists of a 'U-Net' generator (G) and a convolutional neural network (CNN) discriminator (D). We use two different datasets containing visible and thermal images to train our model, and use the structural similarity index measure (SSIM) [2] to measure the performance of our proposed cGAN. Our generated images have good visual effects and reach high SSIM.

First proposed in 2014 [3], GAN has attracted great attention in many fields. A GAN consists of a generator (G) and a discriminator (D). The goal for G is to generate fake samples that are real enough to fool D. For D, its goal is to distinguish real samples from collected databases and fake samples generated by G. By training G and D simultaneously, they can compete with each other and achieve an equilibrium where G can implicitly learn the distribution of the collected databases.

Many image-to-image translation tasks, such as semantic labels to photos and architectural labels to photos, use cGAN to create target images and achieve satisfactory results [4]. For our V2T task, the condition is the input visible image, and our cGAN will generate a corresponding thermal image. Our T2T task is most likely to the face againg task solved by Wang et al. [5] using a cGAN module, an identity-preserved module and an age classifier. For our T2T task, we also use a cGAN module, an identity-preserved module and a temperature classifier to generate fake thermal images in different temperatures.

The paper is organized as follows: Section 2 describes the architecture of cGAN we use. Section 3 details the databases used for training and the pre-processing steps applied to the images. Section 4 summarizes the results of using cGAN approach. Section 5 concludes our paper.

2. PROPOSED METHOD

In our work, we use cGANs to solve both of the image-to-image translation tasks. Our cGAN consists of a generator (G) and a discriminator (D). The structure of G and D for V2T and T2T conversions are slightly different and will be described in Subsection 2.1. For V2T conversion, the input of G is a 3-channel RGB visible image, and the expected output is a thermal image. The input of D is 1 visible image and 1 thermal image, and the output is whether the thermal image is real or fake. For T2T conversion, the input of G is a 1-channel gray-scale thermal image and a target temperature, and the output is the generated 1-channel gray-scale thermal image with the target temperature. The input of D is a 1-channel gray-scale thermal image and a temperature. Only when the thermal image is real and the temperature can represent the input thermal image, the output is real. In other cases, the output is fake. By training G and D together, they can reach an equilibrium with G generating a thermal image that is real enough to fool D.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

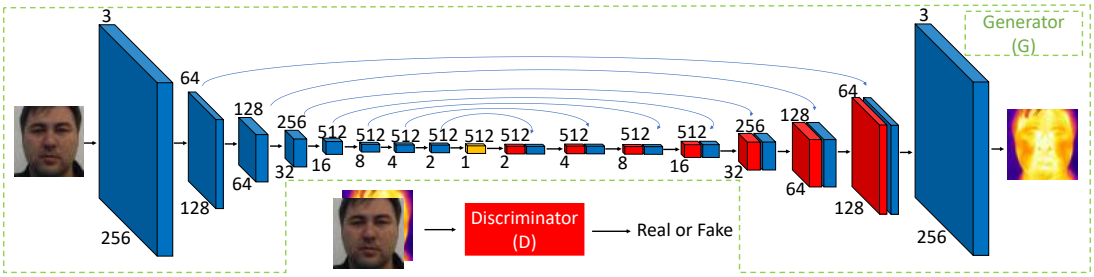


Fig. 1. cGAN network for V2T conversion. The generator (G) is a 'U-Net' with 7 encoding blocks and 7 decoding blocks.

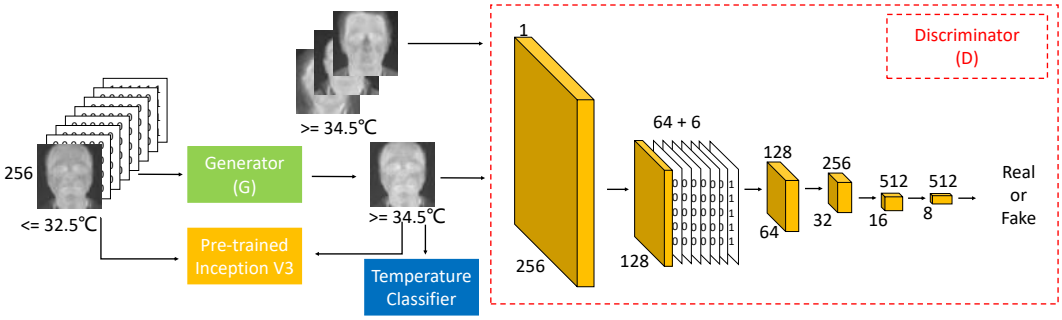


Fig. 2. cGAN network for T2T conversion. The discriminator (D) is a 6-layer CNN.

2.1. Proposed Frameworks

We use 'U-Net' [6] as our generator (G) and a 6-layer conventional neural network as our discriminator (D) for V2T conversion, both using convolution-BatchNorm-LeakyRelu [7].

Figure 1 shows the network structure of our G. Our G consists of 1 input block, 7 encoding blocks, 1 bottleneck, 7 decoding blocks, and 1 output block. Each encoding block will down-sample image size by $1/4$ ($1/2$ of width and $1/2$ of height) of the previous block with $strides = 2$, and each decoding block will up-sample the previous block by 4 times. For the i^{th} decoding block, we add a direct skip from the last i^{th} encoding block, and concatenate the two blocks in the channel before applying the LeakyRelu activation function. The filter size is set to $4 * 4$ for all blocks.

Figure 2 shows the network structure of our D. Our D is a 6-layer CNN, the filter numbers of the first 5 layers are set to 64, 128, 256, 256, 512, respectively, and the stride is set to 2. The last layer is a fully-connected layer, with the activation set as Sigmoid. For T2T conversion, we need to inject temperature information into the networks, so the input layer of G and the second layer of D in T2T is different from V2T, with 6 more channels to incorporate the temperature information.

2.2. Objective Function

The objective function of cGAN for V2T can be written as:

$$\mathcal{L}_G^{V2T}(G,D) = \mathbb{E}_{(x)} \log(1 - D(x, G(x)))$$
$$\mathcal{L}_D^{V2T}(G,D) = -\mathbb{E}_{(x,y)} \log D(x,y) - \mathbb{E}_{(x)} \log(1 - D(x, G(x)))$$

where G tries to minimize $\mathcal{L}_G(G,D)$, and D tries to minimize $\mathcal{L}_D(G,D)$. Additionally, we appended the perceptual loss $\mathcal{L}_{prp}$ used by Jojnson et al. [8]. This consists of the features computed from each single layer of the pre-trained Inception V3 network [9], given by:

$$\mathcal{L}_{prp}^{V2T}(G) = \mathbb{E}_{(x,y)} \sum_{i=1}^{N} \frac{1}{V_i} [||F^{(i)}(y) - F^{(i)}(G(x))||_1]$$

where $F^{(i)}$ denotes the i^{th} layer with V_i activations of the Inception V3 network, and N is the selected number of layers in Inception V3 network. In this experiment, we empirically choose 5 activation layers of Inception V3 network as F to calculate $\mathcal{L}_{prp}$. With $\mathcal{L}_{prp}$, we are able to keep both low-level image characteristic and high-level perceptual information, so that the generated thermal images can keep the identity information as the input image. Combining these losses together, our final objective is expressed as:

$$G_{loss}^{V2T} = \mathcal{L}_G^{V2T}(G,D) + \alpha L_{prp}^{V2T}(G)$$
$$D_{loss}^{V2T} = \mathcal{L}_D^{V2T}(G,D)$$

where α controls the weight of $\mathcal{L}_{prp}^{V2T}$ with respect to $\mathcal{L}_{cGAN}^{V2T}$. Here, we set $\alpha = 10$.

With the temperature information, the objective function of T2T cGAN can be written as:

$$\mathcal{L}_G^{T2T}(G,D) = \mathbb{E}_{(x)} \log(1 - D(G(x,t),t))$$
$$\mathcal{L}_D^{T2T}(G,D) = -\mathbb{E}_{(x)} \log D(x,t) - \mathbb{E}_{(x)} \log(1 - D(G(x,t),t))$$

and the perceptual loss can be written as:

$$\mathcal{L}_{prp}^{T2T}(G) = \mathbb{E}_{(x)} \sum_{i=1}^{N} \frac{1}{V_i} [||F^{(i)}(x) - F^{(i)}(G(x,t))||_1]$$

Apart from this, we use a temperature loss to enforce the generated thermal image fall into the target temperature group. To get the temperature loss, we train a temperature classifier. We divide thermal faces with different temperatures into 6 non-overlapping groups, based on the temperature matrix provided by [10]. Our temperature classifier is adapted from Alexnet [11]. The temperature classifies has the same architecture as Alexnet, except that the last fully-connected layer has only 6 units. The temperature loss is defined as:

$$\mathcal{L}_{temperature}(G) = \mathbb{E}_{(x)} \sigma\left(TC(G(x,t)), t\right)$$

Here $\sigma()$ corresponds to a softmax loss, and $TC()$ is our trained temperature classifier. Through back-propagation, $L_{temperature}$ forces the parameters of G to change and generate faces that fall in the correct temperature group. Combining these losses together, our final objective is expressed as:

$$G_{loss}^{T2T} = \mathcal{L}_{G}^{T2T}(G,D) + \alpha L_{prp}^{T2T}(G) + \beta L_{temperature}$$
$$D_{loss}^{T2T} = \mathcal{L}_{D}^{T2T}(G,D)$$

with α and β both set to 10.

3. EXPERIMENT SETUP

In this paper, we use two databases. The first is the SpeakingFaces Database [12] where subjects are required to read commands, with their voice recorded by a microphone and their thermal faces and visible faces recorded by a FLIR T540 thermal camera and a Logitech C920 Pro HD web-camera. In our work, only the thermal faces in the iron-bow palette and visible face images taken in the front angle from the SpeakingFaces Database are used ($142 * 2 * 8100/9 = 255,600$ thermal-visible image pairs).

The second database is the Carl Database [10], in which visible and gray-scale thermal images containing human faces are collected simultaneously using a thermographic camera TESTO 880-3. The database contains 41 subjects. For each subject, four image acquisition sessions were performed within 2 months, each with 3 different lighting settings (natural, infrared and artificial), and 5 images for each lighting setting ($41 * 4 * 3 * 5 = 2,460$ visible-thermal image pairs).

To train our V2T model, we use the visible-thermal image pairs from the SpeakingFaces Database. We use images collected from subject $1 \sim 100$, all together $100 * 2 * 900 = 180,000$ image pairs to train the model, and use images collected from subject $101 \sim 142$, all together $42 * 2 * 900 = 75,600$ images to test the model. We used SSIM to evaluate the similarity of our generated thermal images and the ground-truth thermal images.

In our experiment, we used images collected from subject $1 \sim 31$, all together $31 * 4 * 3 * 5 = 1,860$ images to train the model, and used images collected from subject $32 \sim 41$, all together 600 images to test the model. We used structural similarity index measure (SSIM) [2] to evaluate the similarity of our generated thermal images and the ground-truth thermal images. We divided $2,460$ thermal face images into 6 groups, based on the face temperature provided in the Carl Database. The temperature range and distribution of these 6 groups are given as follows: ≤ 32.5, $32.5 \sim 33.0$, $33.0 \sim 33.5$, $33.5 \sim 34.0$, $34.0 \sim 34.5$, and ≥ 34.5.

3.1. Alignment, Face Extraction and Resizing

To train the cGAN model, we use $256 * 256$ resolution face image pairs. In Carl Database, the visible and thermal images are not aligned. In our experiment, we used the facial landmark positions manually annotated by Alperen [13] to extract the faces from visible and thermal images, then resizing them into $256 * 256$ resolution. Fig 3 shows an instance of visible-thermal image pair in the original version and after alignment and resizing operation. Based on the 6 facial-landmark-position pairs (blue points), we learn the coordinate mapping from visible image to thermal image. Then we apply a pre-trained face detector `dlib` [14], to extract the face from the visible image (solid green box), and we use the coordinate mapping to map the solid green box in the visible image into the thermal image in order to extract the face from the thermal image (dashed green box). At last, we resized the two boxes into $256 * 256$ resolution (red box) to get our aligned visible-thermal image pair.

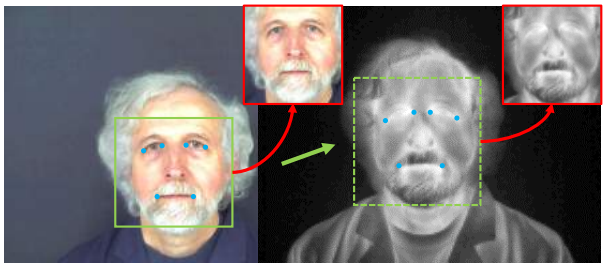


Fig. 3. Alignment and resizing operation of visible and thermal images.

Visible and thermal image pairs in the SpeakingFaces Database are already aligned. We use the pre-trained face detector `dlib` [14] to extract the faces from the visible images, and use the same coordinates to extract faces from thermal images. Then we resize the extracted face images into $256 * 256$ resolution.

4. RESULTS

For the V2T conversion, we present 3 instances of input visible images, ground-truth thermal images, generated thermal images and the corresponding SSIM values in Figure 4. The generated thermal images reach satisfactory visual effects. However, in some parts of the generated images, our model

doesn't perform well. Generated thermal image of the instance in column 1 seems to grow some beard. And for the second instance, our generated thermal image can't reproduce the light shadow in the glasses. We use SSIM to measure the similarity of the generated thermal images and the ground-truth thermal images, and the average SSIM of all images in test set is 0.82 ± 0.06.

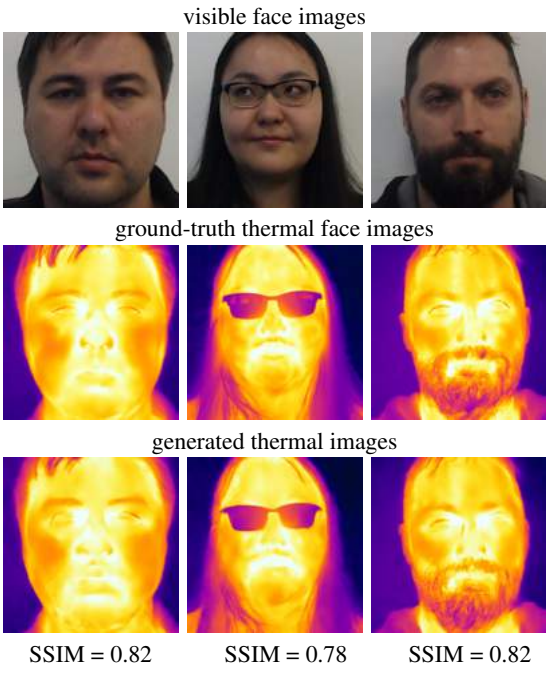


Fig. 4. 3 instances of visible images, generated thermal images and the corresponding SSIM values.

For T2T conversion, we list 3 instances of input thermal images and the different generated thermal images with different target temperatures in Figure 5. And the 3 input thermal images fall in the 3 different temperature groups, from lowest to highest. Looking at Figure 5 from left to right, we see the thermal images become lighter due to they have a higher facial temperature. Ideally, the generated thermal images in the diagonal (red border) should be the same as the input thermal images (blue border). We use SSIM to measure the similarity between the input images and the generated images with the target temperature set to the temperature of the input images, and the average SSIM of the images in the test set is 0.84 ± 0.05. Apart from this, we also use the trained temperature classifier to predict the temperature label of the generated thermal images, and list the confusion matrix in Table 1. We can see that the predicted label is either the expected label or its adjacent labels, and the total accuracy of the prediction is 91.5%.

5. CONCLUSION

In this paper, we solved two image-to-image translation tasks using cGAN. To convert visible images into thermal images, we train our cGAN with cGAN loss and perceptual loss, and the SSIM of the generated thermal images can reach 0.82. To convert thermal images into thermal images with a target temperature, we train our cGAN with cGAN loss, perceptual loss and temperature loss, and the SSIM of the generated thermal images can reach 0.84.

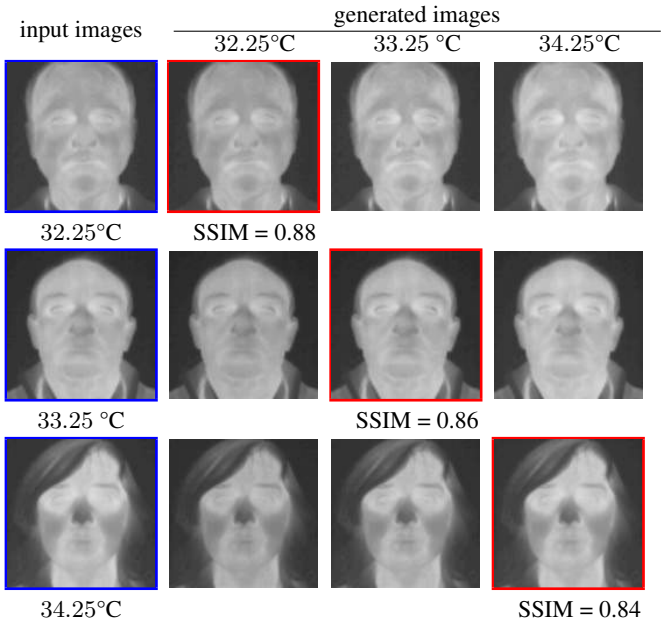


Fig. 5. Generated thermal images in different target temperature range. From left to right are: input thermal image, generated thermal image with different target temperature range. In each row, the input image (blue border) has the same temperature range as the generated image (red border).

Table 1. Confusion Matrix of Generated Thermal Images

Predicted labels (±0.25°C) \ Expected labels (±0.25°C)	32.25	32.75	33.25	33.75	34.25	34.75
32.25	568	25	0	0	0	0
32.75	32	545	25	0	0	0
33.25	0	30	542	28	0	0
33.75	0	0	33	535	31	0
34.25	0	0	0	37	532	27
34.75	0	0	0	0	37	573

6. ACKNOWLEDGE

This Project was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through grant "Biometric-enabled Identity management and Risk Assessment for Smart Cities", and the Mitacs Globalink Graduate Fellowship, and the Department of National Defence's Innovation for Defence Excellence and Security (IDEaS) program, Canada.

7. REFERENCES

[1] Mehdi Mirza and Simon Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[2] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[3] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in *Proceedings of the Advances in neural information processing systems*, 2014, pp. 2672–2680.

[4] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[5] Zongwei Wang, Xu Tang, Weixin Luo, and Shenghua Gao, "Face aging with identity-preserved conditional generative adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7939–7947.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[7] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[8] Justin Johnson, Alexandre Alahi, and Li Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proceedings of the European Conference on Computer Vision*. Springer, 2016, pp. 694–711.

[9] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.

[10] Virginia Espinosa-Duró, Marcos Faundez-Zanuy, and Jiří Mekyska, "A new face database simultaneously acquired in visible, near-infrared and thermal spectrums," *Cognitive Computation*, vol. 5, no. 1, pp. 119–135, 2013.

[11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the Advances in neural information processing systems*, 2012, pp. 1097–1105.

[12] Madina Abdrakhmanova, Askat Kuzdeuov, Sheikh Jarju, Yerbolat Khassanov, Michael Lewis, and Huseyin Atakan Varol, "Speakingfaces: A large-scale multimodal dataset of voice commands with visual and thermal video streams," *arXiv preprint arXiv:2012.02961*, 2020.

[13] Alperen Kantarcı and Hazım Kemal Ekenel, "Thermal to visible face recognition using deep autoencoders," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG)*. IEEE, 2019, pp. 1–5.

[14] Davis E King, "Dlib-ml: A machine learning toolkit," *The Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

BUILDING AND MEASURING TRUST IN HUMAN-MACHINE SYSTEMS

Lida Ghaemi Dizaji, Yaoping Hu

Department of Electrical and Software Engineering
University of Calgary
Calgary, AB, CANADA

ABSTRACT

In human-machine systems (HMS), trust placed by humans on machines is a complex concept and attracts increasingly research efforts. Herein, we reviewed recent studies on building and measuring trust in HMS. The review was based on one comprehensive model of trust – IMPACTS, which has 7 features of intention, measurability, performance, adaptivity, communication, transparency, and security. The review found that, in the past 5 years, HMS fulfill the features of intention, measurability, communication, and transparency. Most of the HMS consider the feature of performance. However, all of the HMS address rarely the feature of adaptivity and neglect the feature of security due to using stand-alone simulations. These findings indicate that future work considering the features of adaptivity and/or security is imperative to foster human trust in HMS.

Index Terms— trust, trust model, measuring trust, human-machine systems

1. INTRODUCTION

Trust placed by humans on machines to undertake designated tasks reflects human willingness to interact with the machines. The higher level of trust is, the more often humans engage the interaction. Thus, human-machine systems (HMS) interconnect humans, machines, and interaction together. Studies have demonstrated that increasing such human interaction in HMS reduces uncertainty and results in an elevated level of trust in HMS [1].

Trust has many application-oriented definitions with multi-dimensional features [2]. A proper model is necessary to evaluate the definitions from various perspectives. The evaluation aids understanding what features are neglected to cause insufficient trust in HMS. Departing from a survey of trust models [3], we focus herein on reviewing studies related to building and measuring trust in HMS. This review is based on the IMPACTS model that is comprehensive with multiple features and specific for assessing trust in HMS [1]. The outcomes of the review suggest future research efforts.

We organize this paper as follows: We first present definitions and modelling of trust and then describe measurement approaches of trust, introducing basic terminologies for the review. After, we set out to report our method and outcomes of the review along with discussing some key observations.

2. DEFINITIONS AND MODELLING OF TRUST

There are various definitions of trust pertinent to HMS. The Merriam-Webster Dictionary lists trust as "assured reliance on the character, ability, strength, or truth of someone or something". The ISO/IEC 25010:2011 standard indicates trust to be "degree to which a user or other stakeholder has confidence that a product or system will behave as intended" [4]. In literature, examples of trust definitions are humans' "willingness towards behavioral dependence" of machines [5] and whether machines' actions and behaviours "correspond to human interest or not" [3].

Three factors affecting trust in HMS are related to humans, machines, and their tasks [6]. The machine-related factor is dominant for building trust, followed by the task-related factor. From a machine perceptive, machines need to be reliable to ensure robust completion of their designated tasks. The tasks must serve the intended purposes of aiding humans to achieve their goals. The human-related factor ranks the least when considering gender and understanding the machines. However, human competence, benevolence, and risk perception play very important roles in building trust placed on the machines [7]. Relying on effective and transparent communication between humans and machines, human trust enables indeed trustworthiness of the machines [8]. Thus, trustworthiness reflects the reliability of the machines from a human perspective and can be absolute (bivalent) and relative (fuzzy) depending on degrees of human confidence.

In general, trust placed by humans on machines is similar to human-human trust [9]. To investigate trust in HMS, there are existing research efforts. Some efforts focus on improving transparent verbal and non-verbal communications [10–12]; while others target identifying task effects on trust under various scenarios [9, 13]. Trust in HMS can be subjective or objective. Subjective trust arises from the perception of individuals, whereas objective trust results from logged data [8].

Despite the efforts, modelling trust has been largely elusive. A recent model – IMPACTS [1] – attempts to be comprehensive, encompassing the above factors and beyond. The

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

model outlines 7 features for HMS to maintain reliability as the result of trust. The features include intention (I), measurability (M), performance (P), adaptivity (A), communication (C) transparency (T), and security (S). In other words, machine behaviors should align with human intentions; the behaviors need to be measurable; the performance of HMS must be reliable and consistent over time, valid as intended and dependable with a low frequency of errors, and predictable to meet human expectations; the HMS should be flexible to re-establish trust when circumstances disrupt the trust; operations and feedback should flow easily between humans and machines; there must be intuitive human-machine interfaces to convey real-time information (such as actions, intentions, decisions, and goals) between humans and machines; and the HMS should behave as designed even under potential attacks. Due to this comprehensiveness, we conducted a review on existing studies of trust in HMS based on the IMPACTS model.

3. MEASURING TRUST

Measuring trust is essential for assessing trust in HMS. Such measurements yield values of some properties related to trust. Since trust has multi-dimensional features according to the IMPACTS model, the values are often qualitative from various perspectives – e.g., predictability, reliability, dependability, and so on [1]. A three-level differentiation – like calibrated trust, over-trust, and distrust – exemplifies such qualitative measurements [14]. The qualitative measurements are easy for humans to understand but difficult for computational analyses. One remedy is quantitative measurements, yielding numeric values based on subjective and/or objective observations. Approaches of the quantitative measurements are questionnaires, behavioural logging, and physiological recording.

Conventionally, questionnaires use numeric scales (e.g., the Likert's scale [15]) to convert human perception of certain trust features into quantitative values. This measurement approach is subjective due to relying on human recalls of the perception and can vary widely in values because of an individual's own perspectives about what to recall. However, the approach is easy to administrate. Many existing research efforts thus utilize questionnaires as a default approach [16–18].

While interacting with machines, the behaviors of humans are indicative of a level of trust placed on the machines. One example is pedestrian behaviors – like willingness to cross a street in front of an automated vehicle, speed of the crossing, head turn for checking the vehicle while crossing, etc. – to indicate certain levels of trust in the vehicle [11]. Another example is engineer behaviors of initiating interactive commands – namely real-time response, simultaneous interaction, and conflict resolution – to ensure the trustworthiness of a collaborative environment for group work [19].

Measuring such human behaviors usually employs data-logging, being objective for analyses. Measuring trust can also derive from physiological signals, such as skin responses and brain activities, associated with human behaviors [5, 20]. As a new approach of objective measurements, physiological recordings demand high skills of mastering recording apparatus and data processing. This demand usually leads to large expenses for the measurements.

Existing research efforts utilize the three approaches either stand-along [15, 20, 21] or certain combinations for measuring trust [5, 22, 23]. The measurements mainly take place within computer-based simulations. Most simulations utilize regular monitors and keyboards/mice to mimic interactive tasks of HMS [16, 24]. Some simulations take the advantages of advanced virtual reality (VR) technology to provide 3D stereoscopic view and interactivity to implement the tasks of HMS [5]. VR simulations trend forwards for measuring trust in HMS, deserving a special consideration in our review.

4. METHOD

4.1. Search Strategy

We conducted the review using the PRISMA method [25]. To identify relevant studies, we searched electronic databases including ScienceDirect, ACM Digital Library, IEEE Xplore, and Google Scholar. Keywords for the search were "trust", "human machine system", and "measure". The search spanned over 5 years (2016 – 2021) to cover the latest studies on HMS. The last date of the search was on 04 March 2021.

4.2. Eligibility Criteria

The selection of eligible studies among the search outcomes applied inclusion and exclusion criteria. The inclusion criteria were two folds: (a) the title, abstract, or content of a study had all search keywords, covering investigation and measurement of trust in HMS; and (b) the study was published in an English, peer-reviewed journal. Due to the page limit, we excluded studies published as reviews, books, thesis, reports, conference articles, or written in other languages.

4.3. Selected Studies

The search of the databases yielded a total of 494 records. Analyses of the records' titles and abstracts, with the elimination of duplicates, produced 121 records for further considerations. Among the considered records, the inclusion and exclusion criteria resulted in 24 studies eligible for review.

5. RESULTS AND DISCUSSION

5.1. Quantitative Analysis

The review applied quantitative and qualitative analyses to the 24 studies. A quantitative analysis of the studies revealed an increasing trend of publications in the past 3 years, as depicted in Fig. 1. This trend shows more research efforts being

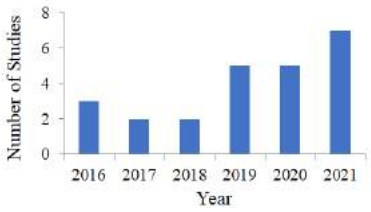


Fig. 1: Temporal distribution of the studies under review.

Fig. 2: Word cloud from the titles of the studies under review.

devoted to trustworthy HMS. A word cloud (generated from https://www.displayr.com/), as presented in Fig. 2, illustrates the most repetition of certain terms in the titles of the studies. The terms are "automation", "human", "trust", "machine", "effects", "reliability", "performance", and "measure". The term "automation" includes "autonomous" and "automated", while the term "measure" covers "measures" and "measured". All the terms concur with the relevance of the studies. Combinations of the terms are thus pertinent to keywords search of building and measuring trust in HMS.

Figure 3 gives a distribution of approaches used for measuring trust among the studies. The approaches are questionnaire (Q), behavioural (B), physiological (P), and their combinations. About 63% of the studies employ mainly questionnaire measurements, 17% use behavioral measurements, 8% apply physiological measurements, and 12% utilize combined measurements. Reflecting in questionnaires, the subjectivity of humans affects mainly the measurement outcomes of the studies. Remedies to the subjectivity are objective approaches such as behavioural, physiological, and combined measurements. Although being a small portion, the combined measurements could offer vital correspondences between subjective and objective measurements.

5.2. Qualitative Analysis

Table 1 summarizes the outcomes of qualitative analysis. For the studies under review, the analysis examined their task, use of VR simulation, supporting each feature of the IMPACTS model, and measurement approach. A ✓ mark indicates that a study supports a particular feature of the model. In contrast, a ✗ mark means that the study has no support for a certain feature. A * mark shows a feature addressed partially by the study, whereas a - mark depicts a feature ignored by the study.

In Table 1, each of the studies has one of the following tasks: operating machines, classifying objects, and managing things. Interestingly, only 4 of the 24 studies use VR simu-

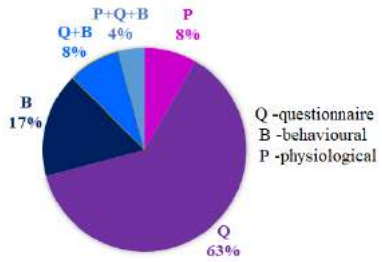


Fig. 3: Percentages of measurements used in the studies.

lations for executing their tasks. In terms of the IMPACTS model, all of the studies take account the features of intention, measurement, communication, and transparency. It is not surprising that the studies address the feature of measurement, because of the inclusion criteria. Most of the studies consider the feature of performance but neglect the feature of security, and address rarely the feature of adaptivity. The feature of security plays a role in HMS, which are vulnerable to be hacked through the Internet. However, most of the studies utilize stand-along simulations for their tasks and, in turn, have no need of considering the feature of security. Nevertheless, the feature of security is important in building trust and deserves the attention of consideration. This importance elevates higher in HMS like autonomous cars and autonomous weapons, which could endanger the lives of humans. Regarding the feature of adaptivity, most of the studies inconsiderate a requirement of rebuilding trust in HMS under unfortunate situations, where humans lose their trust in the machines of HMS. The requirement demands the HMS to be adaptive to handle such situations and to regain human trust. Thus, future research efforts could address this feature for prompting trust.

5.3. Discussion

The above analyses shed some insights on how current HMS prompt human trust. Despite different tasks involved in the studies, there are two common observations. Firstly, human trust in HMS has recently attracted more and more research efforts, even though no consistently agreeable approaches exist for measuring trust. Inappropriate approaches could mislead research efforts, whereas proper and objective approaches would yield accurate and reliable measurement outcomes. It is thus crucial to establish a proper framework of objective measurements of trust. A key step towards the framework is to attest correspondences between objective measurements and subjective perception of trust. Indeed, some studies have attempted the attesting through combining questionnaire, behavioural, and physiological measurements.

Secondly, the IMPACTS model is comprehensive and can serve for building and investigating trust in HMS. Evidently, all studies under review have addressed 4∼5 of the model's7 features without specifying intents to fulfill the model. On one hand, trustworthy HMS should fulfill all 7 features of the model. On another hand, each feature could vary its contri-

Table 1: Summary of the studies under review.

Study	Task	VR Simulation	I	M	P	A	C	T	S	Measurement
[26]	Operating machines - car	Yes	✓	✓	✓	✗	✓	✓	-	Q
[24]	Operating machines - car	No	✓	✓	✓	✗	✓	✓	-	Q
[20]	Operating machines - car	No	✓	✓	✓	✓	✓	✓	-	P
[21]	Operating machines - car	Yes	✓	✓	✓	✓	✓	✓	-	B
[15]	Operating machines - car	Yes	✓	✓	*	✗	✓	✓	-	Q
[16]	Operating machines - aircraft	No	✓	✓	✓	✗	✓	✓	-	Q
[17]	Operating machines - aircraft	No	✓	✓	*	✗	✓	✓	-	Q
[27]	Operating machines - aircraft	No	✓	✓	✓	✗	✓	✓	-	Q
[18]	Operating machines - aircraft	No	✓	✓	✓	✗	✓	✓	-	Q
[28]	Operating machines - aircraft	No	✓	✓	✓	✗	✓	✓	-	Q
[29]	Operating machines - spacecraft	No	✓	✓	✓	✗	✓	✓	-	Q
[22]	Operating machines - unmanned vehicles	No	✓	✓	✓	✓	✓	✓	-	Q+B
[30]	Classifying objects - product	No	✓	✓	*	✗	✓	✓	-	B
[5]	Classifying objects - shapes/words	Yes	✓	✓	✓	✗	✓	✓	-	P+Q+B
[31]	Classifying objects - emails	No	✓	✓	✓	✓	✓	✓	-	Q
[32]	Classifying objects - buildings	No	✓	✓	✓	✓	✓	✓	-	B
[33]	Classifying objects - buildings	No	✓	✓	✓	✓	✓	✓	-	B
[34]	Classifying objects - baggage	No	✓	✓	✓	✗	✓	✓	-	Q
[35]	Classifying objects - baggage	No	✓	✓	✓	✗	✓	✓	-	Q
[23]	Classifying objects - cyber-attacks	No	✓	✓	✓	✗	✓	✓	✓	Q+B
[36]	Managing things - chat-bots	No	✓	✓	✓	✗	✓	✓	-	Q
[37]	Managing things - resources	No	✓	✓	✓	✓	✓	✓	-	B
[38]	Managing things - inventory	No	✓	✓	✓	✓	✓	✓	-	Q
[39]	Managing things - spacecraft air quality	No	✓	✓	✓	✗	✓	✓	-	Q

[Note: VR - virtual reality; IMPACTS: I - intention, M- measurability, P - performance, A - adaptivity, C - communication, T - transparency, S - security; Measurement: P - physiological, B - behavioural, Q - questionnaire]

bution weight to building trust depending on the tasks of the HMS. Factors affecting such weights could be: proximity between a human operator and HMS, the state of the operator's vulnerability,the control of the operator over HMS, and so on. For example, the features of communication and transparency need to have higher weights in a wheelchair HMS than in a gaming counterpart, as human operators of the wheelchair HMS are physically vulnerable and require aiding apparatus to ensure their control over the HMS. Thus, the trustworthiness of HMS depends on their tasks. Increasing communication between humans and machines could elevate trust but might burden the cognitive workload of humans. Since trust is inversely related to cognitive workload [5], this increase must be carefully managed to balance trust and cognitive work-load. Moreover, enhancing transparency could improve trust by alleviating cognitive workload. Such interconnections among certain features of the IMPACTS model exemplify a complexity to achieve trust. Certainly, further efforts are needed to address such interconnections for creating trustworthy HMS. The efforts could take advantages of VR technology to flexibly simulate HMS tasks.

6. CONCLUSION

A unique definition of human trust in HMS remains underexplored, despite many efforts. We reviewed recent studies on building and measuring trust based on the IMPACTS model. The review indicated a positive trend in the studies and suggested future work to address feature interconnections of the model for catalyzing trustworthy HMS.

7. ACKNOWLEDGEMENTS

The authors acknowledge this work is supported by the Department of National Defence's Innovation for Defence Excellence and Security (IDEaS) program, Canada and an NSERC Alliance – Alberta Innovates Advance grant, Canada.

8. REFERENCES

[1] M. Hou et al., "Impacts: a trust model for human-autonomy teaming," *Hum. Intell. Syst. Integr.*, pp. 1–19, 2021.

[2] B. Hernandez-Ortega, "The role of post-use trust in the acceptance of a technology: Drivers and consequences," *Technovation*, vol. 31, pp. 523–538, 2011.

[3] Z. R. Khavas et al., "Modeling trust in human-robot interaction: A survey," in *Proc. ICSR*, 2020, pp. 529–541.

[4] "Systems and software engineering—systems and software quality requirements and evaluation (square)—system and software quality models," ISO/IEC 25010, 2011.

[5] k. Gupta et al., "Measuring human trust in a virtual assistant using physiological sensing in virtual reality," in *Proc. IEEEVR*, 2020, pp. 756–765.

[6] P. A. Hancock et al., "A meta-analysis of factors affecting trust in human-robot interaction," *Hum. Fac.*, vol. 53, pp. 517–527, 2011.

[7] S. Gulati et al., "Design, development and evaluation of a human-computer trust scale," *Behav. & Info. Tech.*, vol. 38, pp. 1004–1015, 2019.

[8] Y. Wang et al., "A tripartite theory of trustworthiness for autonomous systems," in *Proc. IEEE SMC*, 2020, pp. 3375–3380.

[9] J. Y. Jian et al., "Foundations for an empirically determined scale of trust in automated systems," *Int. J. Cogn. Ergon.*, vol. 4, pp. 53–71, 2000.

[10] N. Woodward et al., "Exploring interaction design considerations for trustworthy language-capable robotic wheelchairs in virtual reality," in *Int. Wksh. VAM-HRI*, 2020, 7 pages, vol. 3.

[11] S. Jayaraman et al., "Pedestrian trust in automated vehicles: Role of traffic signal and av driving behavior," *Front. Robot. AI*, vol. 6, pp. 117, 2019.

[12] K. Saleh et al., "Towards trusted autonomous vehicles from vulnerable road users perspective," in *IEEE SysCon*, 2017, pp. 1–7.

[13] S. Shahrdar et al., "Human trust measurement using an immersive virtual reality autonomous vehicle simulator," in *Proc. AIES*, 2019, pp. 515–520.

[14] J. Joe et al., "Identifying requirements for effective human-automation teamwork," in *Proc. 12th PSAM-370-1, 7 pages*, 2014.

[15] S.A. Kaye et al., "Young drivers' takeover time in a conditional automated vehicle: The effects of hand-held mobile phone use and future intentions to use automated vehicles," *Transport. Res. F-TRAF*, vol. 78, pp. 16–29, 2021.

[16] K. Le Goff et al., "Agency modulates interactions with automation technologies," *Ergon.*, vol. 61, pp. 1282–1297, 2018.

[17] J. C. Ferraro et al., "Effects of automation reliability on error detection and attention to auditory stimuli in a multi-tasking environment," *Apld. Ergon.*, vol. 91, pp. 103303, 2021.

[18] N. D. Karpinsky et al., "Automation trust and attention allocation in multitasking workspace," *Apld. Ergon.*, vol. 70, pp. 194–201, 2018.

[19] A. Erfanian et al., "Framework of multiuser satisfaction for assessing interaction models within collaborative virtual environments," *IEEE Trans. Hum. Mach. Syst.*, vol. 47, pp. 1052–1065, 2017.

[20] W.L. Hu et al., "Real-time sensing of trust in human-machine interactions," *IFAC-PapersOnLine*, vol. 49, pp. 48–53, 2016.

[21] M. A. Benloucif et al., "Online adaptation of the level of haptic authority in a lane keeping system considering the driver's state," *Transport. Res. F-TRAF*, vol. 61, pp. 107–119, 2019.

[22] K. Stowers et al., "The impact of agent transparency on human performance," *IEEE Trans. Hum. Mach. Syst.*, vol. 50, pp. 245–253, 2020.

[23] C. Gay et al., "Operator suspicion and human-machine team performance under mission scenarios of unmanned ground vehicle operation," *IEEE Access*, vol. 7, pp. 36371–36379, 2019.

[24] L. Pipkorn et al., "Driver conflict response during supervised automation: Do hands on wheel matter?," *Transport. Res. F-TRAF*, vol. 76, pp. 14–25, 2021.

[25] D. Moher et al., "Preferred reporting items for systematic reviews and meta-analyses: the prisma statement," *Int. J. Surg*, vol. 8, pp. 336–341, 2010.

[26] A. Voinescu et al., "The utility of psychological measures in evaluating perceived usability of automated vehicle interfaces–a study with older adults," *Transport. Res. F-TRAF*, vol. 72, pp. 244–263, 2020.

[27] J. Weyer, "Confidence in hybrid collaboration. an empirical investigation of pilots' attitudes towards advanced automated aircraft," *Safety Sci.*, vol. 89, pp. 167–179, 2016.

[28] P. M. A. De Jong et al., "Time and energy management during approach: A human-in-the-loop study," *J. Aircraft*, vol. 54, pp. 177–189, 2017.

[29] A. Drepper et al., "Multi-layer human-robot teaming: From earth to space," *IFAC-PapersOnLine*, vol. 52, pp. 121–126, 2019.

[30] N. Douer et al., "Theoretical, measured and subjective responsibility in aided decision making," *arXiv preprint arXiv:1904.13086*, 2019.

[31] J. Chen et al., "Automation error type and methods of communicating automation reliability affect trust and performance: An empirical study in the cyber domain," *IEEE Trans. Hum. Mach. Syst.*, pp. 1–11, 2021.

[32] K. Akash et al., "Improving human-machine collaboration through transparency-based feedback–part ii: Control design and synthesis," *IFAC-PapersOnLine*, vol. 51, pp. 322–328, 2019.

[33] K. Akash et al., "Human trust-based feedback control: Dynamically varying automation transparency to optimize human-machine interactions," *IEEE Cont. Syst. Mag.*, vol. 40, pp. 98–116, 2020.

[34] A. Chavaillaz et al., "Some cues are more equal than others: Cue plausibility for false alarms in baggage screening," *Apld. Ergon.*, vol. 82, pp. 102916, 2020.

[35] T. Rieger et al., "Visual search behavior and performance in luggage screening: effects of time pressure, automation aid, and target expectancy," *CRPI*, vol. 6, pp. 1–12, 2021.

[36] J. Balakrishnan et al., "Role of cognitive absorption in building user trust and experience," *Psy. & Mktg.*, vol. 38, pp. 643–668, 2021.

[37] S. Uslu et al., "A trustworthy human–machine framework for collective decision making in food–energy–water management: The role of trust sensitivity," *Knowl. Based Syst.*, vol. 213, pp. 106683, 2021.

[38] L. H. Barg-Walkow et al., "The effect of incorrect reliability information on expectations, perceptions, and use of automation," *Hum. Fac.*, vol. 58, pp. 242–260, 2016.

[39] J. Sauer et al., "The use of adaptable automation: Effects of extended skill lay-off and changes in system reliability," *Apld. Ergon.*, vol. 58, pp. 471–481, 2017.

QUALITY ASSURANCE CHALLENGES FOR MACHINE LEARNING SOFTWARE APPLICATIONS DURING SOFTWARE DEVELOPMENT LIFE CYCLE PHASES

Md Abdullah Al Alamin and Gias Uddin, DISA Lab, University of Calgary

ABSTRACT

In the past decades, the revolutionary advances of Machine Learning (ML) have shown a rapid adoption of ML models into software systems of diverse types. Such Machine Learning Software Applications (MLSAs) are gaining importance in our daily lives. As such, the Quality Assurance (QA) of MLSAs is of paramount importance. Several research efforts are dedicated to determining the specific challenges we can face while adopting ML models into software systems. However, we are aware of no research that offered a holistic view of the distribution of those ML quality assurance challenges across the various phases of software development life cycles (SDLC). This paper conducts an in-depth literature review of a large volume of research papers that focused on the quality assurance of ML models. We developed a taxonomy of MLSA quality assurance issues by mapping the various ML adoption challenges across different phases of SDLC. We provide recommendations and research opportunities to improve SDLC practices based on the taxonomy. This mapping can help prioritize quality assurance efforts of MLSAs where the adoption of ML models can be considered crucial.

Index Terms— Machine Learning Software Application (MLSA), SDLC, ML Pipeline, Challenge, Quality Assurance

1. INTRODUCTION

Each disrupting change in software development required the software industry to evolve and adapt a novel strategy. The latest trend is the widespread interest in the adoption of ML (Machine Learning) capabilities into large scale Machine Learning software applications (MLSAs) [1]. The market of MLSA and Artificial Intelligence (AI) is expected to grow at a compound annual growth rate of 29.7% worldwide. However, quality assurance challenges for MLSAs are hard to address, leading to deadly errors. For example, recently, a Uber self-driving car ran into pedestrians because the sensor could not detect those. There is now a growing concern on the quality assurance of safety-critical ML applications like self-driving cars [2], healthcare, financial institutions, etc.

Software quality assurance is a systematic approach to detect defects and it can improve the reliability and adaptation of MLSA [3]. However, ML models' stochastic nature and their dependency on data introduce diverse novel challenges like testing non-determinism in the MLSA outputs [3]. While deep learning (DL) models have revolutionalized ML models' performance across domains, DL models often behave like a black-box, which makes it difficult to be evaluated and explained in safety-critical applications [4].

In traditional software development, first we gather requirements. We then design, develop, test, deploy and maintain the application. For ML systems, we still need to scope out the goal of the application, but instead of designing the the algorithm we let the ML model learn the desired logic from data [1]. Such observations lead to the question of whether and how ML models can be adopted without disrupting the software development life cycle (SDLC) of the MLSAs. Ideally, *ML workflow/pipeline* and *SDLC* phases should go hand in hand to ensure proper quality assurance. However, as we noted above, such expectations can be unrealistic due to the inherent differences in how ML models are designed and how traditional software applications are developed. We thus need a holistic view of the diverse ML adoption challenges we encounter during the SDLC phases of an MLSA development. Such insights can be used to improve the ML adoption pipeline and the SDLC phases of MLSA development.

In recent years, significant research efforts are devoted to understanding the diverse quality assurance challenges while adopting ML models into software systems [5, 6, 7]. However, we are not aware of any research that specifically focused on mapping the QA challenges across SDLC phases. In this paper, by conducting an in-depth literature review of the challenges of ML adoption and by determining how such challenges may permeate, we have produced a taxonomy of quality assurance challenges that practitioners face during the adoption of ML models into the diverse SDLC phases. We present the taxonomy and describe it with examples. We conclude by offering recommendations for research opportunities that lie ahead to ensure the quality assurance of MLSAs.

2. BACKGROUND AND RELATED WORK

We introduce two concepts (SDLC and ML pipeline stages) on which this paper builds on. We then briefly discuss related work. More related work are discussed in Section 3.

• **Background.** SDLC methodology is embedded in traditional application development that consists of Requirement analysis, Designing & planning, Implementation, Quality as-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

surance, Deployment, Maintenance phases. Our interest is to study the QA challenges faced during different SDLC phases of MLSA. Machine learning workflow is a little bit different from traditional application development. It begins with system requirement analysis to data collection and processing, feature engineering and training, evaluation of model performance and then model deployment and monitoring. Figure 1 shows a typical ML model development stages/workflows [1]. It is an iterative process based on model's performance.

• **Related Work.** There are quite a few numbers of research that focus on the current challenges of machine learning in software engineering[5, 7, 6], empirical studies on best practices of integrating AI capabilities[1], developers survey on challenges faced during different SDLC[8]. There are also quite a few research on the quality assurance of ML models like reliability, transparency, trustworthiness, etc. [9, 10, 11, 6, 10, 12]. Zhang et al.[3] provides a comprehensive survey on ML testing considering 138 related papers on four aspects of MLSA testing such as what properties to test, what ML components to test, testing workflows and test scenario of different type of ML application. Shafiq et al.[13] conducts a bibliometric study and provides an MLSA taxonomy of the adaptation of ML techniques across different SDLC phase. It focuses on the research attention at different SDLC of MLSA. Lack of modularity in ML models compared to traditional software can make those hard to debug [3]. Test oracle problem, due to insufficient specification and data dependency, can introduce a wide range of challenges during software testing [14]. While the above papers offer information of ML adoption challenges, none unlike us, has focused on providing a holistic view of the ML adoption challenges within the standard development life cycle of MLSAs.

3. QUALITY ASSURANCE CHALLENGES IN MLSA

In Figure 2, we present a taxonomy of quality assurance challenges for MLSA development. The challenges are derived from an in-depth literature review of quality assurance challenges for ML adoption into software systems. We group the challenges under six SDLC phases: (1) Requirement analysis, (2) design and planning, (3) implementation, (4) testing, (5) deployment, and (6) maintenance. For each challenge, we also show the specific stages in ML pipeline that are impacted/consulted to address the challenge. The ML pipeline stages are derived from Amershi et al. [1].

In total, we found 31 challenges that we group into three higher categories: (1) *Data*. It contains QA challenges related to data collection, cleaning, labelling, management, (2) *Practice*. It contains QA challenges that is faced in practice by SE teams. (3) *Standard*. This category of QA challenges represent the challenges due to the lack of Standard specification or guidelines. A challenge can be observed in multiple SDLC phases. For example, lack of ML standards and metrics can be a problem during MLSA requirement collection as well as during MLSA implementation phase. However, depending on the types of SDLC, the same challenge can be observed in varying formats. For example, while lack of standards in ML model evaluation may result in ambiguities during MLSA requirement collection, it can also affect the development team during their analysis of whether an ML model is good enough with regards to the stated requirements and designs. We now discuss the challenges per SDLC phases below.

• **SDLC Phase 1. MLSA requirement Collection.** Two challenges arise due to the lack of clarity in *Standard* specification: (1) *Ambiguous requirement (ML stage = R, i.e., Requirements)*: MLSAs can have *ambiguous specification*, as it is difficult to define the expected behaviour, which can change due to data [8, 15]. (2) *Lack of standards in reporting & metrics (R)*: MLSAs do not have a standard reporting specification like other industries (transportation, aviation).

• **SDLC Phase 2. Design and planning.** There are four challenges under three categories: Data, Practice, and Standard. *Data* category has one challenge. (1) *Privacy concerns (R)*: Many companies can not use raw user-data because of terms and service agreement with the users, organizational legal and ethical constrains. Inability to understand the internal representation of the model might cause privacy breach [16]. Federated learning [17] has the potential to deliver these features properly [18]. *Practice* category has two challenges. (1) *Bias and Lack of Fairness (R)*: An MLSA can be biased if the training is biased [19]. (2) *Communication gap between ML and non-ML teams/practitioners (ML Stage = A, i.e., all)*: Qualitative and quantitative studies suggest that it is important to adopt best of SE and ML research world such as sharing data [8], complexity reduction, deletion of features, improving reproducibility [20]. There is one challenge faced due to lack of *Standard* specifications: *Inadequate standard regulation/compliance (R)*: During training we need to consider user data privacy and confidentiality. Organizations need to comply with user data protection acts HIPPA[21], GDPR [22].

• **SDLC Phase 3. MLSA Implementation.** There are eight QA challenges. The *Data* category has one challenge: *Inadequate/Inefficient data management and versioning support (for ML stage = DP, i.e., Data Processing)*. There is a lack of tools to manage different data attributes (e.g., data freshness). The *Practice* category has six challenges: (1) *Lack of modularity & customizability (ML stage = F, i.e., Feature Engineering)*: Due to lack of customizability, it requires a significant effort to reuse a model for a different task or handle different input format. (2) *Longer/impractical evaluation process & time of ML models (ML Stage = T, i.e., Testing)*: The performance of a model can not be evaluated until training on the whole dataset is finished, which is computationally expensive and time-consuming. (3) *Non-deterministic outcome due to data dependency (A)*: ML models are inherently dependent on training data. However, real-world data distribution may be different, and the model might misbehave. (4) *Non-*

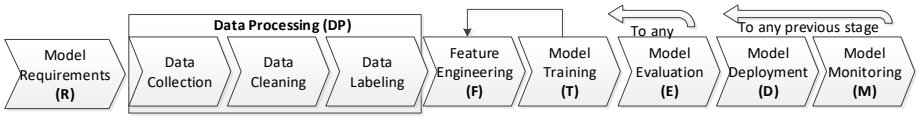


Fig. 1. A typical iterative pipeline (i.e., workflow) to prepare an machine learning (ML) model for a system [1]

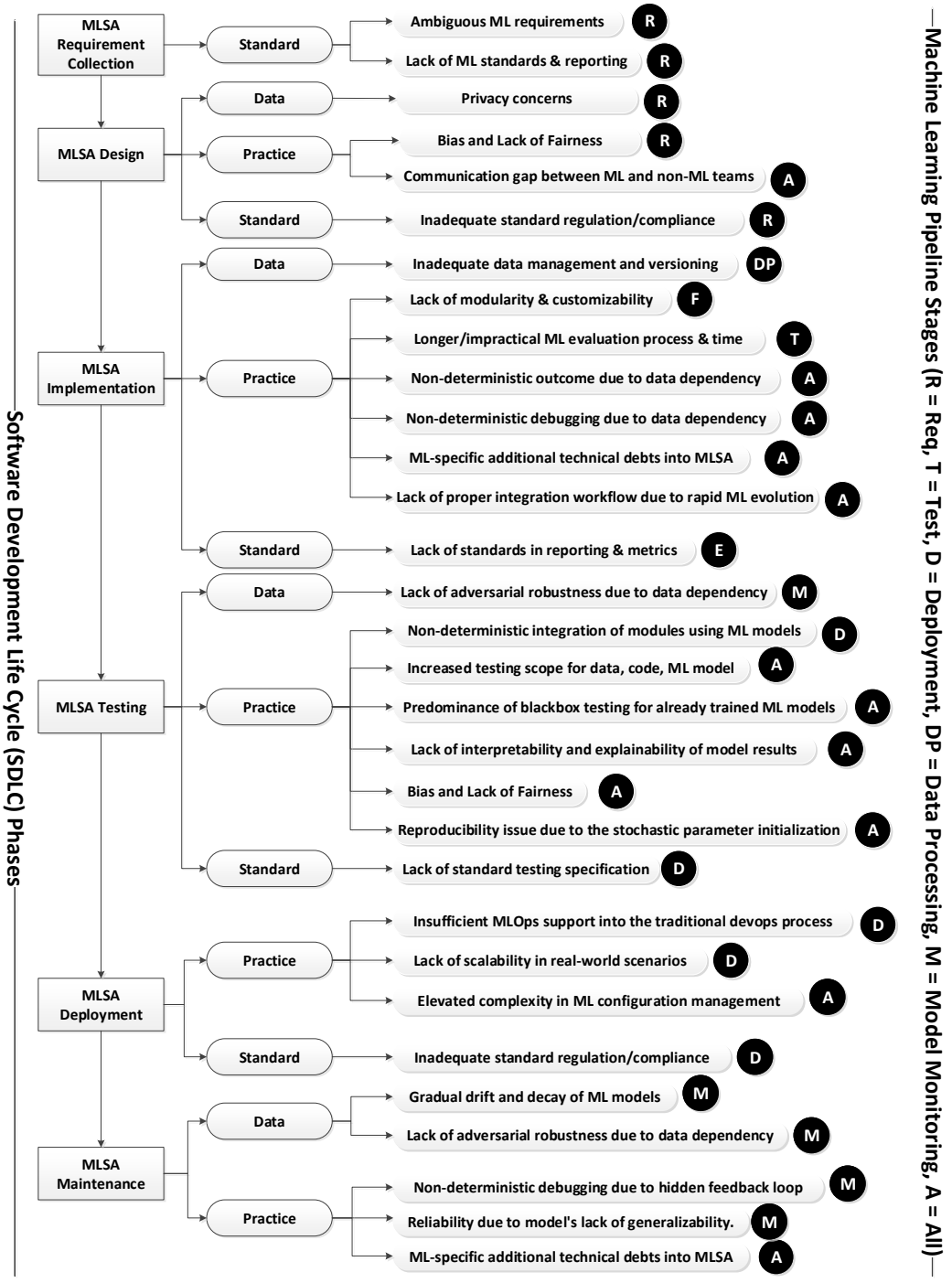


Fig. 2. A taxonomy of MLSA quality assurance challenges across different phases of SDLC

deterministic debugging due to data dependency (A): The debugging strategy of traditional systems is not quite applicable for many ML-systems: bugs might be in the code as well as in the data. (5) *Lack of well-defined workflow for integration due to rapid ML evolution (A)*: There is a lack of ML-specific process management tools, or specification to estimate time and resources. (6) *Presence of additional technical debt due to ML adoption into software (A)*: A mature ML systems can be of 95% of glue code that mainly connects different library and packages [20, 23]. There is 1 QA challenge under *Standard* category: *Lack of standards in reporting & metrics (E)*. There is a lack of standards to compare the performance of different models, guideline to identify & report fault [24].

• **SDLC Phase 4. MLSA Testing.** We find eight challenges. *Data* category has one challenge: *Lack of adversarial robustness due to data dependency (ML Stage = D, i.e., Deployment)*. Adversarial example refers to a small malicious modification of the input that causes the model to erroneous output [25]. The *Practice* category has six challenges: (1) *Non-deterministic integration of ML models (ML stage = M, i.e., monitoring)*: All the external systems that provide input or consumes the output of the ML model should be explicitly monitored because these dependencies can affect the ability to release updates[23, 20]. (2) *Increased testing scope due to the needs for both data and code testing (A)*: QA for ML systems are difficult as they are inherently non-deterministic and they self-learn [1]. (3) *Predominance of black-box-type testing for already trained ML models (A)*: Various phases of a typical MLSA contain highly coupled components. Faults occurred in one component of this pipeline may propagate to other phases and thus hard to detect and fix [26, 27]. (4) *Lack of interpretability and explainability of model results (A)* is prevalent when features are too abstract for human/tester to understand [28, 29]. (5) *Bias and Lack of Fairness (A)*: Model's fairness and biases should be tested in both algorithms and data. (6) *Difficulty in reproducibility due to the stochastic nature of parameter initialization (A)*: Many ML-libraries use random initialization of initial states which makes it hard to reproduce issues. *Standard* category has one challenge: *Lack of standard testing specification (D)*. A model's performance is tested as a whole rather than for a specific input, making it challenging to design a test oracle [14].

• **SDLC Phase 5. MLSA Deployment** There are four QA challenges in this phase and 3 under *Practice* category: (1) *Insufficient MLOps support into the traditional DevOps process (D)*: There is a lack of end to end deployment pipeline support such as advanced logging, automated rollback, security [30]. (2) *Lack of scalability in real-world scenarios (D)*: Many ML models perform well in a research setting but are not computationally scalable for large scale deployment. (3) *Elevated complexity in ML configuration management (A)*: Rapid experimentation requires to track code, data, parameters, hyperparameters to compare the trade-off of different algorithms, model architecture. There is one QA challenges under *Standard* category: *Inadequate standard regulation/compliance (D)*. In order to ensure public trust the model needs to be continually checked to determine whether it is following the user data protection acts (HIPPA[21], GDPR[22]).

• **SDLC Phase 6. MLSA Maintenance.** The maintenance of MLSA over time is expensive and challenging[20]. There are 5 QA challenges. *Data* category has two challenges: (1) *Gradual drift and decay of ML models (M)*: ML systems have a predictable performance degradation over time if the models are not updated with new data. (2) *Lack of adversarial robustness due to data dependency (M)*: The deployed ML-application should perform reasonably in the real world and prevent adversarial attacks[31, 32]. The *Practice* category has three challenges: (1) *Non-deterministic debugging due to hidden feedback loop (M)*: An ML application learns from data and hence some data-driven feedback could be unseen or remain hidden [20]. (2) *Reliability monitoring due to the model's lack of generalizability (M)*: A deployed ML model requires constant monitoring to assess the output against new input and new corner cases. (3) *Presence of additional technical debt due to ML adoption into software (A)*: Many undeclared external systems can consume an ML model's output and cause "visibility debt" [33].

4. RECOMMENDATIONS AND CONCLUSIONS

We conclude by summarizing the research opportunities to address the MLSA quality assurance challenges we categorized across the six SDLC in Figure 2. We find that MLSAs and traditional software applications share some common QA challenges like insufficient specifications and privacy. However, as the performance and robustness of MLSAs depend on the quality of the training dataset, the scope and the severity of these QA challenges differ like non-deterministic outcome, implicit biases in data, lack of interpretability, gradual performance decay, hidden feedback loops pose some of the unique challenges for MLSAs, and so on. To handle ambiguity and lack of standards in *MLSA requirement analysis*, research can focus on developing necessary specifications, tools, and metrics for MLSA requirement analysis and third-party verification [34]. For *design and planning challenges of MLSA*, research can improve data privacy [18, 17, 35] and develop specifications and tools [36] to improve model bias and fairness [19]. For *implementation-specific challenges*, research can focus on developing debugging tools [37], transfer learning [38], model modularization [39], and interpretable AI [28, 29]. To handle the *challenges of testing of MLSA*, more research is needed on bug analysis[3], regression testing, reinforced learning [3], novel adversarial attacks [40, 41, 42, 43] and defence techniques [11, 44]. To address *MLSA deployment and maintenance* challenges, future research can focus on developing configuration management and reliability monitoring specification and tools [31, 32, 9].

5. REFERENCES

[1] Amershi et al., "Software engineering for machine learning: A case study," in *proc ICSE-SEIP*. IEEE, 2019, pp. 291–300.

[2] Chen et al., "Deepdriving: Learning affordance for direct perception in autonomous driving," in *proc ICCV*, 2015, pp. 2722–2730.

[3] Zhang et al., "Machine learning testing: Survey, landscapes and horizons," *In TSE*, 2020.

[4] Willers et al., "Safety concerns and mitigation approaches regarding the use of deep learning in safety-critical perception tasks," in *International Conference on Computer Safety, Reliability, and Security*. Springer, 2020, pp. 336–350.

[5] Devanbu et al., "Deep learning & software engineering: State of research and future directions," *arXiv preprint arXiv:2009.08525*, 2020.

[6] Santhanam et al.r, "Engineering reliable deep learning systems," *arXiv preprint arXiv:1910.12582*, 2019.

[7] Feldt et al., "Ways of applying artificial intelligence in software engineering," in *proc RAISE*. IEEE, 2018, pp. 35–41.

[8] Wan et al., "How does machine learning change software development practices?," *In TSE*, 2019.

[9] Saria et al., "Tutorial: safe and reliable machine learning," *arXiv preprint arXiv:1904.07204*, 2019.

[10] Roscher et al., "Explainable machine learning for scientific insights and discoveries," *IEEE Access*, vol. 8, pp. 42200–42216, 2020.

[11] Goodfellow et al., "Explaining and harnessing adversarial examples," *arXiv preprint arXiv:1412.6572*, 2014.

[12] Qiu et al., "Review of artificial intelligence adversarial attack and defense technologies," *Applied Sciences*, vol. 9, no. 5, pp. 909, 2019.

[13] Shafiq et al., "Machine learning for software engineering: A systematic mapping," *arXiv preprint arXiv:2005.13299*, 2020.

[14] Barr et al., "The oracle problem in software testing: A survey," *IEEE transactions on software engineering*, vol. 41, no. 5, pp. 507–525, 2014.

[15] Finkelstein et al., ""fairness analysis" in requirements assignments," in *2008 16th IEEE International Requirements Engineering Conference*. IEEE, 2008, pp. 115–124.

[16] Davide Castelvecchi, "Can we open the black box of ai?," *Nature News*, vol. 538, no. 7623, pp. 20, 2016.

[17] McMahan et al., "Communication-efficient learning of deep networks from decentralized data," in *AISTATS*. PMLR, 2017, pp. 1273–1282.

[18] Li et al., "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[19] Mehrabi et al., "A survey on bias and fairness in machine learning," *arXiv preprint arXiv:1908.09635*, 2019.

[20] Sculley et al., "Hidden technical debt in machine learning systems," in *Advances in neural information processing systems*, 2015, pp. 2503–2511.

[21] Annas et al., "Hipaa regulations-a new era of medical-record privacy?," *New England Journal of Medicine*, vol. 348, no. 15, pp. 1486–1490, 2003.

[22] Voigt et al., "The eu general data protection regulation (gdpr)," *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, vol. 10, pp. 3152676, 2017.

[23] Arpteg et al., "Software engineering challenges of deep learning," in *proc of Software Engineering and Advanced Applications (SEAA)*. IEEE, 2018, pp. 50–59.

[24] Devanbu et al., "Deep learning & software engineering: State of research and future directions," *arXiv preprint arXiv:2009.08525*, 2020.

[25] Madry et al., "Towards deep learning models resistant to adversarial attacks," *arXiv preprint arXiv:1706.06083*, 2017.

[26] Sculley et al., "Machine learning: The high interest credit card of technical debt," 2014.

[27] Felderer et al., "Quality assurance for ai-based systems: Overview and challenges," *preprint arXiv:2102.05351*, 2021.

[28] Zachary C Lipton, "The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery.," *Queue*, vol. 16, no. 3, pp. 31–57, 2018.

[29] Ribeiro et al., "Why should i trust you? explaining the predictions of any classifier," in *proc ACM KDD*, 2016, pp. 1135–1144.

[30] Paleyes et al., "Challenges in deploying machine learning: a survey of case studies," *preprint arXiv:2011.09926*, 2020.

[31] Schulam et al., "Can you trust this prediction? auditing pointwise reliability after learning," in *proc of Artificial Intelligence and Statistics*. PMLR, 2019, pp. 1022–1031.

[32] Jiang et al., "To trust or not to trust a classifier.," in *proc NeurIPS*, 2018, pp. 5546–5557.

[33] Morgenthaler et al., "Searching for build debt: Experiences managing technical debt at google," in *Workshop on Managing Technical Debt (MTD)*. IEEE, 2012, pp. 1–6.

[34] Vogelsang et al., "Requirements engineering for machine learning: Perspectives from data scientists," in *International Requirements Engineering Conference Workshops (REW)*. IEEE, 2019, pp. 245–251.

[35] Bonawitz et al., "Towards federated learning at scale: System design," *arXiv preprint arXiv:1902.01046*, 2019.

[36] Galhotra et al., "Fairness testing: testing software for discrimination," in *proc ESEC/FSE*, 2017, pp. 498–510.

[37] Ma et al., "Mode: automated neural network model debugging via state differential analysis and input selection," in *proc ESEC/FSE*, 2018, pp. 175–186.

[38] Pan et al., "A survey on transfer learning," *In TKDE*, vol. 22, no. 10, pp. 1345–1359, 2009.

[39] Pan et al., "On decomposing a deep neural network into modules," in *proc ESEC/FSE*, 2020, pp. 889–900.

[40] Dong et al., "Boosting adversarial attacks with momentum," in *proc IEEE conference on computer vision and pattern recognition*, 2018, pp. 9185–9193.

[41] Carlini et al., "Towards evaluating the robustness of neural networks," in *symposium on security and privacy (sp)*. IEEE, 2017, pp. 39–57.

[42] Szegedy et al., "Intriguing properties of neural networks," *arXiv preprint arXiv:1312.6199*, 2013.

[43] Mei et al., "Using machine teaching to identify optimal training-set attacks on machine learners," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2015, vol. 29.

[44] Zhang et al., "Defense against adversarial attacks using feature scattering-based adversarial training," *preprint arXiv:1907.10764*, 2019.

AN OPEN SOURCE MOTION PLANNING FRAMEWORK FOR AUTONOMOUS MINIMALLY INVASIVE SURGICAL ROBOTS

Aleks Attanasio, Nils Marahrens*, Bruno Scaglioni and Pietro Valdastri*

STORM Lab, University of Leeds, United Kingdom

ABSTRACT

Planning and execution of autonomous tasks in minimally invasive surgical robotic are significantly more complex with respect to generic manipulators. Narrow abdominal cavities and limited entry points restrain the use of external vision systems and specialized kinematics prevent the straightforward use of standard planning algorithms. In this work, we present a novel implementation of a motion planning framework for minimally invasive surgical robots, composed of two subsystems: An arm-camera registration method only requiring the endoscopic camera and a graspable device, compatible with a 12mm trocar port, and a specialized trajectory planning algorithm, designed to generate smooth, non straight trajectories. The approach is tested on a DaVinci Research Kit obtaining an accuracy of 2.71 ± 0.89 cm in the arm-camera registration and of 1.30 ± 0.39 cm during trajectory execution. The code is organised into STORM Motion Library (STOR-MoLib), an open source library, publicly available for the research community.

Index Terms— Da Vinci Research Kit (dVRK), Trajectory planning, ROS

1. INTRODUCTION

Trajectory planning lies at the heart of most robotic manipulation tasks and is crucial to enable high levels of autonomy [1]. While tasks usually define a set of different poses to be achieved, how the robot should move in between these poses is often left to motion planning algorithms. Common motion planners integrate a plethora of robot models, but surgical minimally invasive surgical systems are not well represented. This may attributed to their complex kinematic structures, often including parallel chains that are not supported by most inverse kinematics solvers and can be numerically challenging. Moreover, the software frameworks used to control surgical robots such as the Collaborative Robot Toolkit (CRTK) [2] and the DaVinci Research Kit (dVRK) [3] only provide the ability to reach a final pose with zero velocity, thus not supporting the execution of complex trajectories.

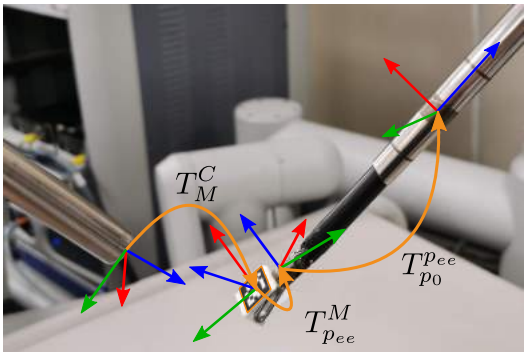


Fig. 1. Transformations of the different frames considered for the registration of the arm to the camera frame.

In the particular case of the dVRK, one of the most popular surgical robotics research platform [4], a point to point trajectory in the joint space is generated from the current end effector pose to the goal by means of the Reflexxes RML II [5] library. The resulting trajectory might be optimized in joint space but is generally neither smooth nor optimal in cartesian space. The available literature on motion planning for surgical robots is scarce. In [6] the problem is addressed for the dVRK platform using the MoveIt![7] motion platform. However, the extended abstract is silent on how the problem of parallel kinematics is solved, nor is their code publicly available to the community. Recent works have focused on employing machine learning techniques, such as Pyramid Stereo Matching Network (PSMNet) [8] and reinforcement learning [9]. While these methods show impressive results on specific tasks, they are not generally applicable and easily adaptable. Moreover, they are highly dependent on large amounts of labeled data, obtained via computationally and time-intensive simulations. Another common problem limiting the development of autonomous tasks in MIS robotics platforms is the co-registration between the camera and the robotic arms, since the two subsystems are usually connected to different bases. This issue is commonly solved for generic manipulators using external optical trackers [10]. This approach has been adopted for surgical robots [8, 11] by attaching markers on the tip of

* these two authors contributed equally. Research reported in this article was supported by Intuitive Surgical Inc. under the Technology Research Grants program 2019, by the Royal Society, by the Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/R045291/1, and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 818045).

the surgical instruments. Although accurate, this method requires the use of an external camera, which is a major limitation in a small and delicate environment such as the abdominal cavity, and is prone to inaccuracies due to the presence blood or debris in the surgical scene. In this work, we: (1) Present a software framework aimed at solving the problem of co-registration for robotic platforms specific to MIS, focused on the ease of use and the feasibility of the application in a clinical environment. (2) Present an approach to the planning and execution of complex trajectories on surgical robots, integrated with ROS and easily adaptable to any platform. (3) Provide public and documented code in a web repository to benefit the surgical robotics research community.

2. CO-REGISTRATION ALGORITHM

This section describes the approach adopted to determine the transformation between the endoscopic camera and the surgical instrument held by the robot. This step is crucial to plan and execute autonomous tasks based on visual servoing in scenarios where the endoscope and the robotic arm do not share the same reference frame. This is the case with robots such as the dVRK, the Raven [12] and modular robots like CMR Versyus or Medtronic's Hugo RAS. The goal is to compute the transformation from the camera frame to the origin of the robotic arm. This can be solved by evaluating a sequence of transformations that start from the pose of the robot end-effector with respect to the camera. In robots equipped with cameras, this can be achieved by adopting a computer vision algorithm to detect one or more visual markers mounted on the end-effector. To this end, we adopt the ArUco markers [13] and mount them on a custom 3D printed pick-up device, designed to be held by standard surgical instruments and be inserted through standard 12mm trocar ports. Once the pick-up device with ArUco marker is grasped by the robotic instrument (Fenestrated Bipolar Forceps), exposed to the camera and recognized by the vision algorithm, the transformation T_C^{p0} between the PSM's base frame T_{p0} and the endoscope's base frame T_C is calculated as follows:

$$T_C^{p0} = T_C^M T_M^{p_{ee}} T_{p_{ee}}^{p_0} \quad (1)$$

where T_C^M is the transformation between camera and a visual marker held by the end-effector, $T_M^{p_{ee}}$ is the transformation between the marker and the end-effector reference frame, and finally $T_{p_{ee}}^{p_0}$ is the pose of the end-effector with respect to the robot base frame. The transformations are shown in Figure 1 on a DaVinci Patient Side Manipulator (PSM), in which the base frame is placed in the remote centre of motion, on the trocar. Assuming that $T_{p_{ee}}^{p_0}$ can be extracted from the robot kinematics and that $T_M^{p_{ee}}$ is known by design of the marker holder, T_C^M can be estimated by using the endoscope in conjunction with software packages like tuw_marker_detection [14]. Finally, the transformation $T_M^{p_{ee}}$ is applied to align the marker frame with the tool tip frame of the robot. To increase robustness of the results, we combine both detected transformations from the left and right endoscopic camera and average the results over 100 frames, each 100ms apart.

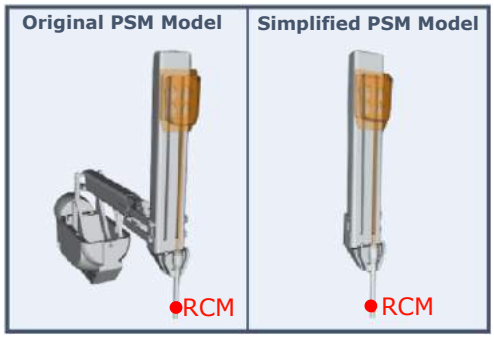


Fig. 2. Original PSM model and the simplified model used in this work. In our simplified model, the base of the robot is omitted, thus removing the parallel kinematic chain and allowing the usage of the MoveIt! package without any loss of generality in the trajectory planning.

3. TRAJECTORY PLANNING

The co-registration algorithm enables to evaluate and control the position of the robot end-effector in the camera workspace. This feature facilitates the definition of points of interest based on computer vision or deep-learning algorithms and to relate them to the position of the end-effector. In many autonomous tasks, it is required to generate a trajectory based on the points identified in this step, and to execute it smoothly. One goal of this paper is to provide a framework for planning and smoothing of the trajectory dedicated of surgical robotic tools. For this purpose, the MoveIt! [15] framework has been used, due to the wide adoption in the research community. MoveIt! is based on the widely used Open Motion Planning Library (OMPL) [16] that includes state-of-the-art algorithms for trajectory planning, manipulation and navigation and is integrated into ROS [17]. In order to plan a trajectory for a specific robot, and therefore produce a feasible trajectory in joint and Cartesian spaces, MoveIt! gathers information about the robot layout from two files: the Unified Robot Description Format file (URDF), used in the ROS ecosystem to define robots kinematics, and the Semantic Robot Description Format file (SRDF), which includes additional information to the URDF such as default robot configuration and collision checking. The trajectory planning is carried out in four steps: (1) The robot URDF and SRDF are loaded onto Moveit!. (2) The robot starting position, way-points and goal of the trajectory are defined. (3) The MoveIt! function computeCartesianPath() is used to evaluate a sequence of points on straight lines from the starting position, through the way-points, to the final goal. (4) The Stochastic Trajectory Optimization for Motion Planning

(STOMP) [18] is used to plan trajectory using the previously generated points as seeds and produce the final trajectory, represented as a set of points in the 3D workspace. STOMP is adopted for its capability of avoiding local minima while allowing a faster convergence to the solution if compared to other planners such as Covariant Hamiltonian Optimization for Motion Planning (CHOMP) [19]. Additionally, given its stochastic nature, the STOMP planner can generate a smooth path even in the presence of obstacles.

A C++ library, STORM Motion Library (STOR-MoLib) is developed to provide the code to the community. The library requires minimal user input and can be utilized by means of the following methods: *compileMotionPlanRequest(waypoints_constraint, trajectory_seed)* and *transformTrajectory(trajectory, base_frame)*. The first populates the MoveIt! motion request constraining the passage through the desired way-points. The trajectory seeds are the output of the computeCartesianPath function included in MoveIt!. The second function transforms the trajectory points from the robot frame to the user-defined base frame, in our case the camera frame. The MoveIt! motion request is then solved by the STOMP Planner which returns a smoothed trajectory.

4. EXPERIMENTAL VALIDATION

The validation of our approach is composed of two steps: the evaluation of the accuracy for the camera-arm registration and the assessment of the trajectories planning and execution. Although the application of the framework could be generalized to any robot, in this work we focus on the dVRK due to its ubiquity and the availability of an open source simulation software, thus circumventing the need for a physical platform, to replicate the results described here. In particular, we adopt a subset of the full DaVinci system composed of one PSM and one stereoscopic endoscope mounted on an independent base. A Linux (Ubuntu 18.04) machine equipped with an Intel Xeon Gold 6140 (2.30GHz) CPU, an Nvidia Quadro 5000 RTX GPU and 128 GB DDR4 2666MHz RAM was adopted to carry out the planning. While the use of a specific robot is transparent to the co-registration algorithm, the trajectory planning depends on the features of each robotic arm through the URDF and SRDF files. Initially, the PSM description files provided with the dVRK library [3] are used. However, the PSM adopts a parallel mechanism to ensure a fixed remote centre of motion. This type of kinematics is not supported in MoveIt!. In order to overcome this issue, a modified version of the PSM excluding the parallel link is developed (Figure 2). Despite the different physical layout, the kinematics of the robot is correctly reproduced by maintaining the Remote Centre of Mass fixed and eliminating the parallel link and the preceding links in the kinematic chain.

To quantify the registration error, a 3D-printed calibration body attachable to the endoscope's tip was designed. The calibration body contains nine landmark points (p_C^1 - p_C^9) with known distance with respect to the camera's base frame T^C (Figure 3a). By touching the landmarks with the tip of the surgical instrument, we acquired the location of these positions in the PSM's base frame T_{p0}.By performing several registrations ($n = 5$) and averaging the position of each of the nine points over all runs we obtain p_{p0}^1 - p_{p0}^9. With a confidence interval of 0.0734 mm ($c = 0.95$), we assume the robot's positional accuracy to be fairly high and consistent compared to the camera. In order to assess the accuracy of the co-registration approach on our surgical setup, five registrations are performed using the ArUco marker with differing tool positions and thus different placements of the marker with respect to the camera. With the acquired transformations $T_C^{p_0}$ from the visual marker registrations, we transform the points p_C^1 - p_C^9 on the calibration body from the camera's base frame T_C to the PSM's base frame T_{p0} and calculate the euclidean distance to the respective points obtained via landmark registration. Our results indicate a mean positional error of 2.71 ± 0.89 cm ($c = 0.95$) over all registered points and registration runs compared to the position obtained via the camera calibration body. We believe the main source of inaccuracy to be the camera distortion. Despite a thorough calibration, the fish-eye lenses of the endoscope produce a significant distortion that negatively affects the accuracy of the marker detection, particular when the marker is not place directly at the center of the image. Additionally, the small distance between the two cameras limits the usage of further information from the 3D scene via stereo matching or similar techniques.

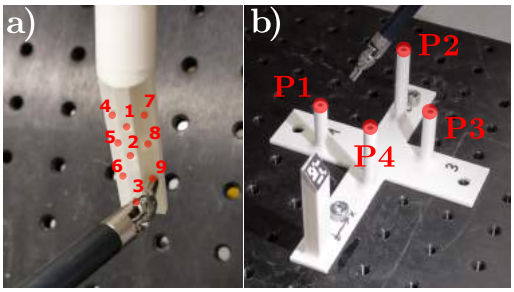


Fig. 3. 3D-printed rigid body used for the validation of the marker-based co-registration (a). 3D-printed rigid body used to validate the precision during the trajectory execution (b). A marker has been attached to the body to allow the registration of the points via the camera.

In order to evaluate the accuracy of the trajectory planning and execution, a 3D-printed reference body with four vertical pegs was designed. The tip of each peg represents either a way-point or the goal of the trajectory (Figure 3b). The reference body also integrates an ArUco marker, added to obtain a transformation from its local reference frame to the camera frame T_C^{RB}. The coordinates of each way-point are transformed into the PSM's base frame T_{p0} by combining the

two previously obtained transformations ($T_{p0}^{RB} = T_{p0}^{C} T_{C}^{RB}$). The planner evaluates a trajectory starting from the current position of the instrument, passing along the way-points and ending in the goal position. Two different trajectory scenarios have been considered with three and four way-points, respectively. Each trajectory has been repeated 8 times and, for each repetition, the surgical instrument was initially manually placed in a varying position around the starting point. Although the planner can consider variable instrument orientations, we maintained a constant, randomly selected, orientation during the whole trajectory.

The planner's output consists of a trajectory defined as an array of joint values, one set for every trajectory point. These are converted to the Cartesian space by means of forward kinematics and eventually organised in a vector of poses sent to the dVRK software. The dVRK only allows a point to point trajectory, constraining the initial and goal velocity to zero. To perform a smooth trajectory, we published the new poses at a rate of 20Hz, sending a new command before the robot had reached the previous goal and thus avoiding the condition of zero velocity. Before executing each trajectory, the position of each way-point with respect to the robot's base frame T_{p0} was collected by manually positioning the surgical instrument (large needle driver) onto a landmark on each peg's tip and recording its position. Figure 4 shows the 8 trajectories for both the three and four point case. The start and end point of the trajectory are represented in blue and green, respectively. The way-points are represented in red. It must be pointed out that the sequence of the way-points is different for the two trajectories. The sequence chosen in the four point case is aimed at demonstrating the ability of the planner to find a solution in the even in the case of more involved trajectories, containing a indirect path with back and forth motion. The evaluation of the trajectories is carried out by considering the minimum distance between the path executed by the robot and each way-point measured before the trajectory execution via the robots tool tip. With this reference, the average error amounts to 1.09 ± 0.59 cm ($c = 0.95$) for the three point and 1.30 ± 0.39 cm ($c = 0.95$) in the four point case.

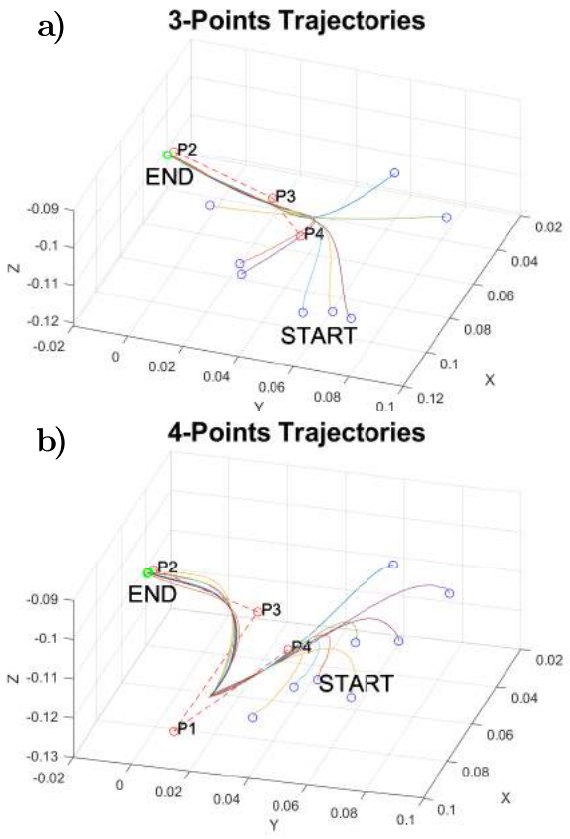


Fig. 4. Repetitions for the trajectory planning and execution for three point (a) and four point (b) case. The initial point is shown in blue, the goal point in green and the way-points in red. The red dashed lines depict the seeds used by the STOMP planner.

5. CONCLUSIONS

In this paper, we presented a comprehensive library to manage the trajectory planning of surgical robots with the specific aim of developing a method that does not require dedicated hardware such as optical trackers or external cameras, thus applicable in the context of minimally invasive surgery. Initially, we presented a method for arm-to-camera registration based on the ArUco markers. We showed the method to be a feasible approach in robotic systems where the arms and the camera do not share the same kinematic base. Subsequently, we demonstrated an approach for planning and executing trajectories based on Moveit! and integrated with ROS. For our evaluation, we applied our framework and approach to the dVRK platform. The registration makes it possible to plan trajectories with respect to the camera frame, thus supporting the execution of vision-based autonomous surgical gestures. Moreover, the registration algorithm can be useful in setups, such as the dVRK, in which teleoperation is challenging due to the lack of a simple built-in co-registration protocol. Although the dVRK Setup Joints controller will be available in the future, not all the research groups have access to the full platform. We believe that this library could significantly benefit the research community. STOR-MoLib code is open source and publicly available [1].

Further development of this library, currently under investigation, include the implementation of a collision avoidance algorithm, useful in collaboration scenarios in which a human operator is controlling one arm, while the other arm is autonomously operated. Other improvements, particularly regarding the registration accuracy, might be obtained by further investigations on the distortion of the cameras' lenses.

[1] https://github.com/Stormlabuk/dvrk_stormolib

6. REFERENCES

[1] Aleks Attanasio, Bruno Scaglioni, Elena De Momi, Paolo Fiorini, and Pietro Valdastri, "Autonomy in surgical robotics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, 2020.

[2] Anton Deguet Peter Kazanzides, "CRTK - Collaborative Toolkit," 2021, [Online; accessed 21-April-2021].

[3] Peter Kazanzides, Zihan Chen, Anton Deguet, Gregory S. Fischer, Russell H. Taylor, and Simon P. Dimaio, "An open-source research kit for the da Vinci® Surgical System," in *IEEE International Conference on Robotics and Automation (ICRA)*. 2014, pp. 6434–6439, IEEE.

[4] Claudia D'Ettorre, Andrea Mariani, Agostino Stilli, Ferdinando Rodriguez y Baena, Pietro Valdastri, Anton Deguet, Peter Kazanzides, Russell H. Taylor, Gregory S. Fischer, Simon P. DiMaio, Arianna Menciassi, and Danail Stoyanov, "Accelerating Surgical Robotics Research: Reviewing 10 Years of Research with the dVRK," apr 2021.

[5] Torsten Kroeger, "Opening the door to new sensor-based robot applications - The reflexxes motion libraries," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 6–9, 2011.

[6] Zhixian Zhang, Adnan Munawar, and Gregory S Fischer, "Implementation of a motion planning framework for the davinci surgical system research kit," in *The Hamlyn Symposium on Medical Robotics*, 2014, p. 43.

[7] Sachin Chitta, Ioan Sucan, and Steve Cousins, "Moveit![ros topics]," *IEEE Robotics & Automation Magazine*, vol. 19, no. 1, pp. 18–19, 2012.

[8] Florian Richter, Shihao Shen, Fei Liu, Jingbin Huang, Emily K Funk, Ryan K Orosco, and Michael C Yip, "Autonomous robotic suction to clear the surgical field for hemostasis using image-based blood flow detection," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1383–1390, 2021.

[9] Florian Richter, Ryan K Orosco, and Michael C Yip, "Open-sourced reinforcement learning environments for surgical robotics," *arXiv preprint arXiv:1903.02090*, 2019.

[10] Christoff M. Heunis, Beatriz Farola Barata, Guilherme Phillips Furtado, and Sarthak Misra, "Collaborative Surgical Robots: Optical Tracking During Endovascular Operations," *IEEE Robotics & Automation Magazine*, vol. 27, no. 3, pp. 29–44, sep 2020.

[11] Caitlin Schneider, Christopher Nguan, Robert Rohling, and Septimiu Salcudean, "Tracked "pick-Up" ultrasound for robot-assisted minimally invasive surgery," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 2, pp. 260–268, 2016.

[12] Blake Hannaford, Jacob Rosen, Diana W Friedman, Hawkeye King, Phillip Roan, Lei Cheng, Daniel Glozman, Ji Ma, Sina Nia Kosari, and Lee White, "Raven-ii: an open platform for surgical robotics research," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 954–959, 2012.

[13] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, jun 2014.

[14] Markus Bader Lukas Pfeifhofer, "ROS package tuw_aruco," 2019, [Online; accessed 14-April-2021].

[15] Sachin Chitta, Ioan Sucan, and Steve Cousins, "MoveIt!," *IEEE Robotics and Automation Magazine*, vol. 19, no. 1, pp. 18–19, 2012.

[16] Ioan A Sucan, Mark Moll, and Lydia E Kavraki, "The open motion planning library," *IEEE Robotics & Automation Magazine*, vol. 19, no. 4, pp. 72–82, 2012.

[17] Hideki Yoshida, Hiroshi Fujimoto, Daisuke Kawano, Yuichi Goto, Misaki Tsuchimoto, and Koji Sato, "Range extension autonomous driving for electric vehicles based on optimal velocity trajectory and driving braking force distribution considering road gradient information," *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*, pp. 4754–4759, 2015.

[18] Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, Peter Pastor, and Stefan Schaal, "STOMP: Stochastic trajectory optimization for motion planning," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 4569–4574, 2011.

[19] Nathan Ratliff, Matt Zucker, J. Andrew Bagnell, and Siddhartha Srinivasa, "CHOMP: Gradient optimization techniques for efficient motion planning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2009, pp. 489–494.

TOWARDS EXPLAINABLE SEMANTIC SEGMENTATION FOR AUTONOMOUS DRIVING SYSTEMS BY MULTI-SCALE VARIATIONAL ATTENTION

Mohanad Abukmeil‡, Angelo Genovese‡, Vincenzo Piuri‡, Francesco Rundo†, and Fabio Scotti‡
‡ Department of Computer Science, Università degli Studi di Milano, Italy {*firstname.lastname*}@unimi.it
† STMicroelectronics, ADG, Central R&D, 95121 Catania (CT), Italy *francesco.rundo*@st.com

ABSTRACT

Explainable autonomous driving systems (EADS) are emerging recently as a combination of explainable artificial intelligence (XAI) and vehicular automation (VA). EADS explains events, ambient environments, and engine operations of an autonomous driving vehicular, and it also delivers explainable results in an orderly manner. Explainable semantic segmentation (ESS) plays an essential role in building EADS, where it offers visual attention that helps the drivers to be aware of the ambient objects irrespective if they are roads, pedestrians, animals, or other objects. In this paper, we propose the first ESS model for EADS based on the variational autoencoder (VAE), and it uses the multiscale second-order derivatives between the latent space and the encoder layers to capture the curvatures of the neurons' responses. Our model is termed as Mgrad$_2$VAE and is bench-marked on the SYNTHIA and A2D2 datasets, where it outperforms the recent models in terms of image segmentation metrics.

Index Terms— Autonomous Driving System, VAE, XAI, ESS.

1. INTRODUCTION

The rapid advancement of artificial intelligence (AI) and machine learning (ML) has lead to the development of AI-powered autonomous systems, which can sense, learn, decide and interact for many different applications including computer vision, natural language processing (NLP), robotics, autonomous driving, and others fields[1, 2]. Moreover, AI-powered autonomous systems are build based on deep learning (DL) models comprising convolutional neural networks (CNN), autoencoders (AEs), generative adversarial networks (GANs), and Bayesian models [3]. However, the effectiveness of many recent models and systems are limited due to the scarcity of explainability; such an explainability translates the actions and decisions of the learned models to users who operate and develop them. Explainable artificial intelligence (XAI) is a branch of AI aims to explain the behaviors of the ML models [4].

An autonomous driving system (ADS) is referred to any vehicle that can sense the surrounded environment without human control, or with a limited level of supervision. ADS is also able to control engines, visualize objects, detect abnormal actions, drive vehicles, and activate breaks [5]. AI and ML influence ADS by automatically processing data, offering instantaneous recommendations, and recognizing objects; such objects include pedestrians, trees, bicyclers, and other moving and static objects [6, 7]. Explainable autonomous driving systems (EADS) combine XAI and ADS to enhance the vehicular automation (VA), throughout interpreting sensory data, mentoring vehicles behaviors, and semantically segmenting the ambient objects [4]. In this regard, the explainable semantic segmentation (ESS) is a branch of the ML in which each pixel of the segmented object holds a semantic meaning, and can be integrated into the EADS to improve the explainability of the detected objects, and to offer roads conditions conclusion to the drivers [8].

XAI-powered models are associated with unsupervised learning (UL) to visualize the hidden structure of data [9, 1]. AEs are a class of UL methods that are able to generate and visualize data, reduce dimensionality, and perform other ML tasks such as object recognition [10]. AEs comprise classic, de-noising, contractive, sparse, variational-AE (VAE) [10, 11]. Moreover, the success of AEs architectures led to the flourishing of different supervised AEs for structured prediction, i.e., semantic segmentation, such as Seg-net, U-net, and others [12, 13, 14]. Among all AEs, VAE is regulated by the variational inference (VI) to optimize the posterior distribution of large datasets, which leads to a better generalization. The VAEs have been utilized in the ADS in the absence of XAI, where it has been used in the steering control [15], pedestrian prediction in [16], trajectory simulation in [17], and anomaly detection for ADS [18].

The first work towards explaining the VAE behavior is proposed in [19], where it generates visual attention to show how the encoder side behaves. Moreover, the proposed attention map is built by duplicating the last layer of the encoder, thereafter it scales each feature point in the filter channels by a global average pooling of the gradient of the latent space concerning that layer. Factually, the drawback of such attention lies in the unfair scaling, i.e., both related and unrelated feature points are scaled with the same factor. On the other hand, the first work that has been attempted to build the attention of the CNN for ADS is described in [6], where the attention is built by averaging the activations of 100 images; such attention hides the effects of the high and low activations, i.e., approximated attention, and is not stable for time-series segmentation.

To fill the gap of explaining VAEs in the EADS applications, we propose Mgrad$_2$VAE[1], a novel ESS model for EADS applications. Moreover, the Mgrad$_2$VAE utilizes the multiscale second-order derivative between the latent space and each encoder layer, which captures the curvatures of neurons' activations to build the multiscale explainable attention without unfair scaling or averaging the final attention. Therefore, our contribution is twofold: *(i)* introducing a novel ESS model for EADS applications, by using the unsupervised VI and a supervised convolutional AE, and *(ii)* proposing a novel multi-scale gradient attention mapping scheme for ESS to improve EADS applications using the second-order derivative operator. The rest of this paper is organized as follows. Section 2 highlights the VAE and the proposed explanation methodology. Section 3 describes the architecture of Mgrad$_2$VAE. The experimental results are given in Section 4. The conclusion and future works are reported in Section 5.

This work was supported in part by the EC within the H2020 Program under projects MOSAICrOWN and MARSAL and by the Italian Ministry of Research within the PRIN program under project HOPE. We thank the NVIDIA Corporation for the GPU donated.

[1]The source code is available at:
`http://iebil.di.unimi.it/mgradvae/index.htm`

2. VAE AND THE EXPLAINABILITY METHODOLOGY

2.1. VAE

VAEs consist of many different encoding and decoding stages, where each stage represents a different scale of dimensionality that is contracted or expanded by using learning parameters θ (where $\theta = \{W, B\}$, W and B are weights and biases, respectively) [17]. The learning parameters are used to perform many different mapping including convolution, dense multiplication, deconvolution, regularization, etc, by utilizing several sets of representations to capture neurons' activations [20]. Also, for each setting among parameters θ, i.e., after each learning epoch, the gradient of the output is estimated with respect to the input by employing the first-order (1^{st}) partial derivative to optimally reconstruct or generate data.

VAE encompasses two main modules [11]: *(i)* the inference (encoder) module that is used to map an image (or data) $X = \{x_i \,|\, x_i \in \mathbb{R}^D,\, i = 1, \ldots, N\}$, D is the original dimensionality ($D = m \times n \times c$ which indicates rows, columns, and channel depth, respectively), to a latent space $Z = f(X) = \{z_i = f(x_i) \in \mathbb{R}^d, | \, i = 1, \ldots, M\}$. Moreover, the encoder module reduces dimensionality of the data, i.e., $0 < d < D$, and it is used to infer the model likelihood $P(X|\theta)$ [10]. *(ii)* The generation (decoding) module that is utilized to reconstruct the original data $\tilde{X}$ from the latent space Z. For a given data $X \in \mathbb{R}^D$, the encoding module creates a mapping $f : \mathbb{R}^D \rightarrow \mathbb{R}^d$, while the decoding module creates an inverse mapping $g : \mathbb{R}^d \rightarrow \mathbb{R}^D$, which generates an approximation of the data: $\tilde{X} = g(Z; \hat{\theta}_d)$ [21]. Similarly to AEs, VAE is regulated to find the optimal set of parameters $(\hat{\theta}_e, \hat{\theta}_d)$ that achieve a better generalization [14], and to attain the minimum reconstruction loss $\mathbf{L}_{\text{rec}}$:

$$\mathbf{L}_{\text{rec}_{\{\hat{\theta}_e,\, \hat{\theta}_d\}}} = \min \|X - (f \circ g)X\|^2_{\text{Er}} \tag{1}$$

where E_r represents the reconstruction error metric which can be computed by reconstruction cross-entropy, β- divergence, mean square error (MSE), Frobenius norm, or β- divergence [9].

VI is utilized to regulate the VAE, where two different losses are optimized simultaneously for a better generalization [11]. The VI is a Bayesian method that approximates an intractable posterior over a large dataset, throughout approximating the probability densities by optimization. The VAE's encoder approximates the posterior distribution $Q(Z|X)$, which identifies the distributional shape of the latent space Z according to the original data X. Moreover, the VAE is characterized by $Q(Z|X)$ optimization; such an optimization affects the distribution of latent space Z to follow a Gaussian distribution with a definite mean μ (which reflects the Gaussian's center), and standard deviation σ (which reflects the Gaussian's shape).

Practically, the prior distribution of the latent space $P(Z)$ is considered (simply by duplicating the unit Gaussian distribution of th original data manifold $P(X)$); subsequently, the prior $P(Z)$ and the approximated distribution $Q(Z|X)$ are matched by utilizing the KL divergence [22]. The KL divergence is always positive and tends to zero if and only if P and Q are almost equal in the distribution, and it is mathematically defined as $\text{KL}(P\|Q) = \Sigma_x P(x) \log \frac{P(x)}{Q(x)}$. The variational process is known as the reparameterization trick, and it can be obtained by perturbing σ with a small noise ϵ, thereafter directing the optimizer to enforce the AE to reconstruct the data concerning the distribution of X. Moreover, the reparameterization trick augments the generalization, where it produces different distributions to be compared with $P(Z)$ as in duplicating data [11].

Eventually, the VAE optimizes the reconstruction loss $\mathbf{L}_{\text{rec}}$ through minimization in accordance to Eqn. (1), and it is also optimized to minimize the distributional loss of the latent space between $Q(Z|X)$ and $P(Z)$ using $\text{KL}(P\|Q)$, that reflects which extent the reparameterized latent distribution follows a unit Gaussian:

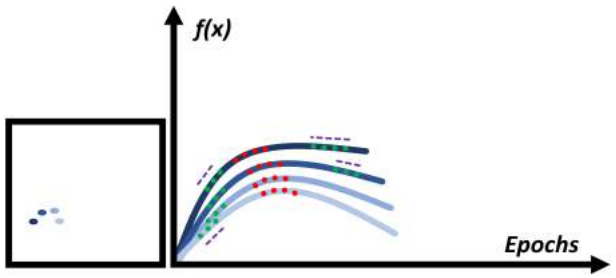


Fig. 1: The neurons activations and gradient over epochs.

$$\mathbf{L}_{\theta_{\text{VAE}}} = \min[\mathbf{L}_{\text{rec}} + \text{KL}(P\|Q)] \tag{2}$$

where $\theta_{\text{VAE}} = \{\hat{\theta}_e, \hat{\theta}_d, \hat{\mu}_X, \hat{\sigma}_X, \hat{\mu_Z}, \hat{\sigma_Z}\}$.

2.2. The explainability methodology

Deep semantic segmentation models comprise many different encoding and decoding blocks to map data from the domain of the original image to the corresponding masks [12, 13]. Moreover, neurons with different parameters (θ_e, θ_d) are employed to optimally fit models, where at each learning epoch the gradient that measures the instantaneous rate of change among the model parameters is measured, by utilizing the first-order partial derivative ∂ between each pixel in the segmented mask with respect to the input image [23].

Considering a VAE with a single encoding layer L_{e1} and a latent layer Z, the first gradient between Z and L_{e1} is estimated according the partial derivative of each neuron activation z_i as $\frac{\partial z_i}{\partial L_{e1}}$. Moreover, if an additional layer L_{e2} lies between L_{e1} and Z, then the chain rule is used as $\frac{\partial z_i}{\partial L_{e1}} = \frac{\partial z_i}{\partial L_{e2}} \frac{\partial L_{e2}}{\partial L_{e1}}$ [24]. The result of all derivations gives the required rate of changes to update θ. Given a period of time, the neuron activations are changing; capturing such variations draws an attention map that gives an insight into how the neurons respond among different inputs, and it is obtained by considering the derivative of the gradient, i.e., 2^{nd} partial derivative $\frac{\partial^2 z_i}{\partial L^2_{e1}}$ [25].

Visually, four pixels of an image with their associated neurons activations are illustrated in Fig. 1, where the activations are given according to the non-linear ReLU functions [26] (the method is valid for other types of activations). The 1^{st} gradient is the slope (magenta dashed lines) at any point in the curves (blue curves), where the derivative of the gradient interprets how the curves are varied during a time (the red and green points). As it is observed from Fig. 1, the gradient of activations can be stationary during a period of the learning time, i.e., the 2^{nd} derivative around the green points is ≈ 0, however, it can vary at a different period of time, i.e., the 2^{nd} derivative around the red points is $>$ or < 0. Accordingly, utilizing the 2^{nd} derivative which measures how the 1^{st} gradient of the activations of the neurons are changing, is able to capture the temporal behaviors of the neurons (as in deriving the acceleration from speed) which reflects the curvatures of learned representations.

Due to the VI, the latent space Z hides many different representations that are generated to regularize the VAE; such representations assist in building ESS attention utilizing the behavior of the neurons' activations. To build a visual attention map, our Mgrad_2VAE aggregates all multiscale derivatives of the gradient of the latent layer Z concerning each encoding layer, which represents a different scale of dimensionality. For a better visual explanation, our proposed attention map is enforced to follow the original mask distribution, by minimizing the reconstruction and KL losses between the reconstructed mask, attention map, and original mask, simultaneously.

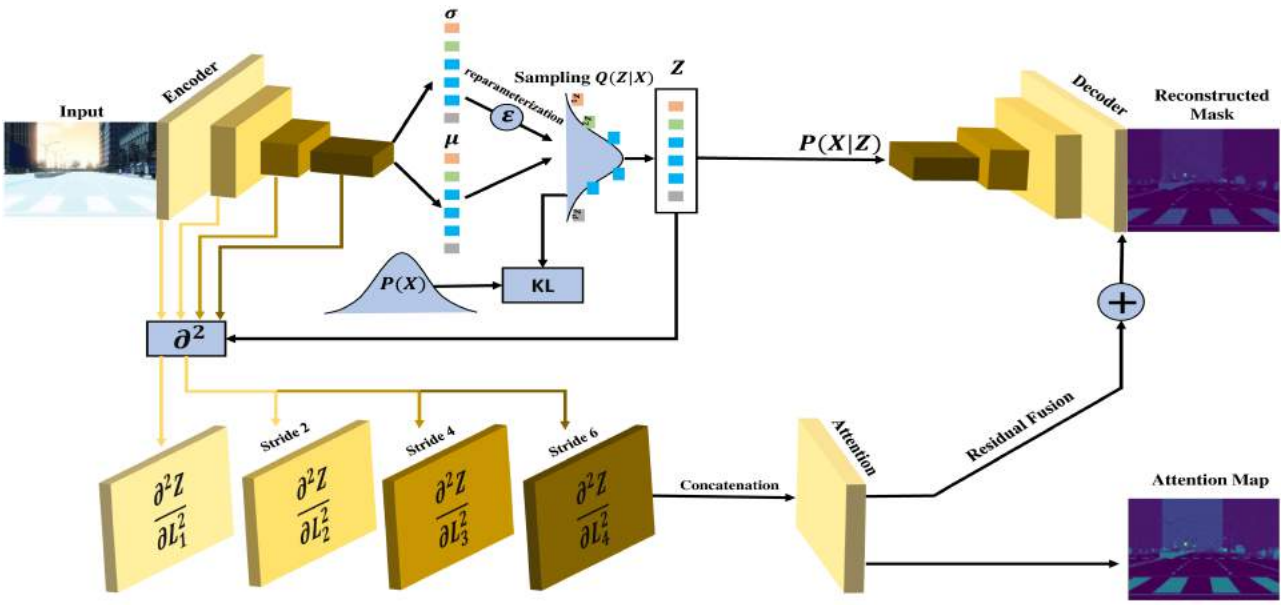


Fig. 2: The Mgrad$_2$VAE block diagram.

3. Mgrad$_2$VAE

Fig. 2 shows our proposed Mgrad$_2$VAE, where it encompasses encoder, decoder, and attention modules. Jointly, the encoder and the decoder include three stages of down-scaling (convolutional neurons with a stride of 2) and up-scaling (de-convolutional neurons with a stride of 2), respectively. Moreover, the Mgrad$_2$VAE visually explains the learned representations utilizing 2^{nd} gradient attention at each encoding scale, i.e., for each encoder's layer there will be a corresponding visual attention map that reflects explainability at that layer, and each attention is enforced to follow the mask distribution to help to contract the representations to the mapped mask.

Moreover, for each layer, the tensor that holds all partial derivatives of the gradient is re-scaled for sake of optimization to match the mask size. Thereafter, all attention maps are aggregated and fused with the $L_{d_{n-1}}$ layer (d_n is the total number of the decoder's layers); such a combination is considered as a novel form of the residual learning [27], which enforces the Mgrad$_2$VAE to learn the residual of mapping between the images and masks by using the 2^{nd} gradient attention. Consequently, besides the explainability of the Mgrad$_2$VAE, it also assists in mask reconstruction by employing the curvatures of activations that are fused to the decoder. Accordingly, the Mgrad$_2$VAE optimizes two losses by using Adam [23] as:

$$\mathbf{L}_{\text{Mgrad}_2\text{VAE}} = \min[\mathbf{L}_{\text{VAE}} + \|X - \theta_{Mgrad}(Z, L_{e_i})\|^2_{\text{Er}}] \quad (3)$$

where the first loss is obtained from the vanilla VAE [11] that is described at Eqn. (2), and the second loss is the reconstruction loss between the original mask and the aggregated attention at the attention module (see Fig. 2). Furthermore, θ_{Mgrad} reflects the 2^{nd} derivative parameters between the latent space Z concerning all encoder layers L_{e_i}, i.e., for each layer, there will be a corresponding tensor of the size of that layer to allocate all partial derivatives, and the final tensor holds the multiscale attention. Additionally, the model is trained to minimize the loss between each mapped image and its corresponding segmentation mask, and it also optimizes the loss between each attention map that is obtained at a different scale with the same mask; such an optimization enforces all encoder layers to contract to the same data, and it compensates the encoding loss that is raised from down-scaling the dimensionality in the depth layers.

4. EXPERIMENTAL RESULTS

To show the performance of our proposed Mgrad$_2$VAE, we used a collection of SYNTHIA [28] and A2D2 [29] datasets. Specifically, 5600 samples are categorized to the corresponding semantic classes that have been employed. Moreover, the dataset partition to the training and testing subsets complies with $75:25$ protocol, i.e., 75% and 25% of the original data size are the training and testing subsets, respectively. For the sake of computation, in the qualitative analysis, the Mgrad$_2$VAE considers an input layer of the size of $128\times256\times3$, where the output layer of the size of $128\times256\times1$. For all experimental works, we consider a minibatch size of 16, and 600 epochs with a learning rate $\eta = 0.001$, where the η is decreased every 100 epoch by a factor of 10^{-2}.

4.1. Qualitative analysis

Our Mgrad$_2$VAE visually explains the learned representations at the neurons activations level through the attention mapping, where it considers the 1^{st} derivative of the gradient (i.e., the 2^{nd} order derivative of neurons activations) between the latent space Z and the encoder layers. For each encoding layer, it produces a tensor to allocate all partial derivatives, and the final attention map can be obtained by concatenating and aggregating (see Fig. 2) all corresponding tensors by using different methods including mean, addition, convolution, etc. Fig. 3 shows the corresponding tensor unfolding (of an image from SYNTHIA dataset) of the attention that is obtained from the last encoding layer L_{e_4}, which represents the last encoding scale as a function of 16 filters depth.

Furthermore, Fig. 4 depicts the final aggregated attention of all encoding layers, where it shows how our model can visually explain the global characteristics of the learned representations at an early stage (L_{e_1}). Moreover, it is also able to show the local characteristics among representations that are captured from the fine-grained features in the depth layers (L_{e_4}).

Fig. 5 shows different examples from the SYNTHIA testing set, ground-truth (GT) masks, reconstructed masks, and the attention maps obtained from our Mgrad$_2$VAE. As it can be noticed from Fig. 4 and Fig. 5, all attention maps which are obtained by our Mgrad$_2$VAE are contracted to the ground truth mask distribution (target domain), and they jointly utilize the multiscale attention mapping (attention at each layer) to build a complimentary map for a

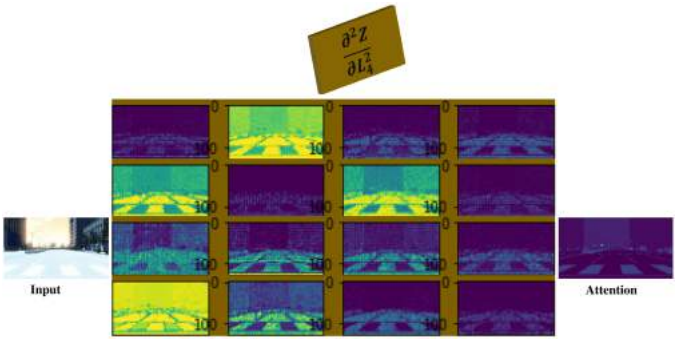


Fig. 3: The 2^{nd} order derivative unfolding of Z with respect to L_{e_4}.

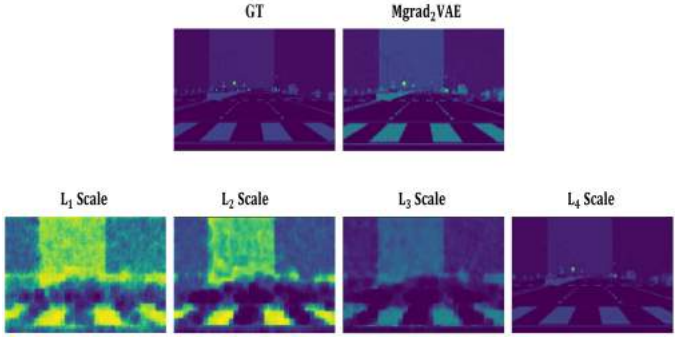


Fig. 4: The multiscale attention of our proposed Mgrad$_2$VAE, where GT represents the ground truth mask.

better visual explainability.

To assess the semantic structure qualitatively, we employ the SSIM index [30] for both datasets. Moreover, we report the semantic similarities between the ground-truth masks, the corresponding reconstructed masks, and the attention maps in Table 1.

SSIM Index	SYNTHIA	A2D2
Reconstructed masks	97.57%	60.38%
Attention maps	96.47%	55.71%

Table 1: SSIM of the reconstructed masks and the attentions.

As it can be noticed from Table 1, the Mgrad$_2$VAE produces an attention map that preserves a similar SSIM index for the reconstructed mask by the decoder, which confirms our methodology and reflects the high quality of the produced attentions.

4.2. Quantitative analysis

Table 2 reports the pixel-wise predictive performance of our proposed Mgrad$_2$VAE, where we consider the average area under the receiver operator characteristic curve (AUC-ROC) index which reflects an aggregated measure of each pixel classification accuracy. Moreover, we consider the same experimental setup that is reported in section 4. For sake of numerical stability, the depth of the output layer has been adapted from $128 \times 256 \times 1$ to $128 \times 256 \times 3$.

As it can be observed from Table 2, our proposed model offers high performance at the pixel-level classification for both the reconstructed masks and attention maps. Moreover, our attention mapping method outperforms the reconstruction obtained from the decoder side in terms of pixel-level classification in the SYNTHIA dataset.

4.3. Recent work comparison

In this section, we compare our proposed Mgrad$_2$VAE model with the recent deep learning models, where we consider the deep VAE [11], and the Xception model [31] that has been built based on the U-net architecture [13] and trained on ImageNet dataset [32]. Moreover, we summarize the AUC-ROC metric between the GT masks and the reconstructed masks among all models in Table 3.

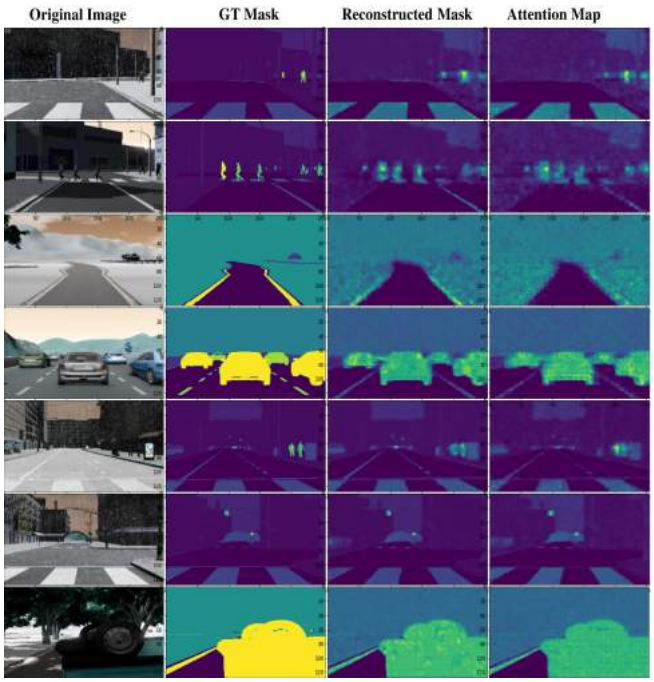


Fig. 5: Examples from the SYNTHIA testing set, where the original images, GT masks, reconstructed masks, and the Mgrad$_2$VAE attention maps are illustrated from left to right, respectively.

AUC-ROC	SYNTHIA	A2D2
Reconstructed masks	81.50%	95.44%
Attention maps	83.20%	95.36%

Table 2: AUC-ROC of the reconstructed masks and the attentions.

AUC-ROC	SYNTHIA	A2D2
Deep VAE [11]	79.60%	94.05%
Xception [31]	67.43%	95.19%
Our Mgrad$_2$VAE reconstruction	81.50%	95.44%
Our Mgrad$_2$VAE attention	83.20%	95.36%

Table 3: AUC-ROC comparison with recent deep models.

As it can be seen from Table 3, our proposed Mgrad$_2$VAE model outperforms all other models in reconstructing masks and attentions. Moreover, although the reconstruction module of our model is typical to the Deep VAE [11], the reconstruction performance of the Mgrad$_2$VAE is better than [11] by 1.90% and 1.39% for the SYNTHIA and A2D2 datasets, respectively, because of the residual fusion between the decoder and attention modules of our model.

5. CONCLUSIONS

We proposed an explainable VAE model termed as the (Mgrad$_2$VAE) to be utilized for XAI and EADS applications. Our model uses the multiscale second-order derivative of the neurons' activations of the latent space concerning all other encoding layers. Moreover, it captures the curvature of the learned representations to offer a better visual explainability of the VAE's behavior through attention mapping. Our proposed model outperforms all related deep segmentation models in the quantitative analysis. In future works, we plan to investigate the XAI in harsh environments and rough weather conditions, where the ambient includes rain, snow, dust, fog, etc.

6. REFERENCES

[1] Mohanad Abukmeil, Stefano Ferrari, Angelo Genovese, Vincenzo Piuri, and Fabio Scotti, "A survey on unsupervised generative models for exploratory data analysis and representation learning," *Acm computing surveys (csur)*, 2021.

[2] Sorin Grigorescu, Bogdan Trasnea, Tiberiu Cocias, and Gigel Macesanu, "A survey of deep learning techniques for autonomous driving," *Journal of Field Robotics*, vol. 37, no. 3, pp. 362–386, 2020.

[3] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García, Sergio Gil-López, Daniel Molina, Richard Benjamins, et al., "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information Fusion*, vol. 58, pp. 82–115, 2020.

[4] David Gunning and David Aha, "Darpa's explainable artificial intelligence (xai) program," *AI Magazine*, vol. 40, no. 2, pp. 44–58, 2019.

[5] Junqing Wei, Jarrod M Snider, Junsung Kim, John M Dolan, Raj Rajkumar, and Bakhtiar Litkouhi, "Towards a viable autonomous driving research platform," in *Proc of Intelligent Vehicles Symposium (IV)*, 2013.

[6] Chenyi Chen, Ari Seff, Alain Kornhauser, and Jianxiong Xiao, "Deepdriving: Learning affordance for direct perception in autonomous driving," in *Proc. of ECCV*, 2015.

[7] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun, "CARLA: An open urban driving simulator," in *Proc. of robot learning*, 2017.

[8] A. Genovese, V. Piuri, F. Rundo, F. Scotti, and C. Spampinato, "Pedestrian/cyclist distance estimation from a single rgb image: A cnn-based semantic segmentation approach," in *Proc. of Industrial Technology (ICIT 2021)*, 2021.

[9] Mohanad Abukmeil, Stefano Ferrari, Angelo Genovese, Vincenzo Piuri, and Fabio Scotti, "On approximating the nonnegative rank: Applications to image reduction," in *Proc. of CIVEMSA*, 2020.

[10] Mohanad Abukmeil, Stefano Ferrari, Angelo Genovese, Vincenzo Piuri, and Fabio Scotti, "Unsupervised learning from limited available data by $\beta-\mathrm{NMF}$ and dual autoencoder," in *Proc. of ICIP*, 2020.

[11] Diederik P. Kingma and Max Welling, "Auto-encoding variational bayes," in *Proc. of ICLR*, 2014.

[12] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos, "Image segmentation using deep learning: A survey," *arXiv preprint arXiv:2001.05566*, 2020.

[13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. of Medical image computing and computer-assisted intervention*, 2015.

[14] Yoshua Bengio, Aaron Courville, and Pascal Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[15] Alexander Amini, Wilko Schwarting, Guy Rosman, Brandon Araki, Sertac Karaman, and Daniela Rus, "Variational ae for end-to-end control of autonomous driving with novelty detection and training de-biasing," in *Proc. of IROS*. IEEE, 2018.

[16] Atanas Poibrenski, Matthias Klusch, Igor Vozniak, and Christian Müller, "M2p3: multimodal multi-pedestrian path prediction by self-driving cars with egocentric vision," in *Proc. of ACM Symposium on Applied Computing*, 2020.

[17] Xinyu Chen, Jiajie Xu, Rui Zhou, Wei Chen, Junhua Fang, and Chengfei Liu, "Trajvae: A variational autoencoder model for trajectory generation," *Neurocomputing*, vol. 428, pp. 332–339, 2021.

[18] Andrea Stocco, Michael Weiss, Marco Calzana, and Paolo Tonella, "Misbehaviour prediction for autonomous driving systems," in *Proc. of ICSE*, 2020.

[19] Wenqian Liu, Runze Li, Meng Zheng, Srikrishna Karanam, Ziyan Wu, Bir Bhanu, Richard J Radke, and Octavia Camps, "Towards visually explaining variational autoencoders," in *Proc. of CVPR*, 2020.

[20] Pierre Baldi, "Autoencoders, unsupervised learning, and deep architectures," in *Proc. of Unsupervised and Transfer Learning workshop*, 2012.

[21] Geoffrey E. Hinton and Ruslan R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[22] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra, "Stochastic backpropagation and approximate inference in deep generative models," in *Proc. of ICML*, 2014.

[23] Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," in *Proc. of ICML*, 2014.

[24] William F Ames, *Numerical methods for partial differential equations*, Academic press, 2014.

[25] Kai Fan, Ziteng Wang, Jeff Beck, James Kwok, and Katherine Heller, "Fast second order stochastic backpropagation for variational inference," in *Proc. of NIPS*, 2015.

[26] Vinod Nair and Geoffrey E Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. of ICML*, 2010.

[27] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proc. of CVPR*, 2016.

[28] Javad Zolfaghari Bengar, Abel Gonzalez-Garcia, Gabriel Villalonga, Bogdan Raducanu, Hamed Habibi Aghdam, Mikhail Mozerov, Antonio M Lopez, and Joost van de Weijer, "Temporal coherence for active learning in videos," in *Proc. of ICCVW*, 2019.

[29] Jakob Geyer, Yohannes Kassahun, Mentar Mahmudi, Xavier Ricou, Rupesh Durgesh, Andrew S Chung, Lorenz Hauswald, Viet Hoang Pham, Maximilian Muhlegg, Sebastian Dorn, et al., "A2d2: Audi autonomous driving dataset," *arXiv preprint arXiv:2004.06320*, 2020.

[30] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[31] Francois Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. of CVPR*, 2017.

[32] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," *Proc. of NIPS*, 2012.

ATTENTIVE AUTOENCODERS FOR IMPROVING VISUAL ANOMALY DETECTION

Ambareesh Ravi, Fakhri Karray

CPAMI, Dept. of Electrical and Computer Engineering, University of Waterloo, ON, CANADA

ABSTRACT

Understanding the notion of normality in visual data is a complex issue in computer vision with plenty of potential applications in several sectors. The immense effort required for optimal design for real-world application of existing methods warrants the need for a generic framework that is efficient, automated and can be momentarily deployed for the operation, reducing the effort expended on model design and hyper-parameter tuning. Hence, we propose a novel, modular and model-agnostic improvement to the conventional AutoEncoder architecture, based on *visual soft-attention for the inputs* to make them robust and readily improve their performance in automated semi-supervised visual anomaly detection tasks, without any extra effort in terms of hyper-parameter tuning. Besides, we discuss the role of attention in AutoEncoders (AE) that can significantly improve learning and the efficacy of the models with detailed experimental results on diverse visual anomaly detection datasets.

Index Terms— Convolutional AutoEncoders, Anomaly Detection, Attention

1. INTRODUCTION

Anomaly Detection (AD) is an important field of machine learning with a significant amount of dedicated research [1]. Anomaly detection is the process of identification of rare, out-of-order and abnormal samples that do not conform with the regular patterns of data. It is now a vast field of research with applications in critical areas such as medical diagnosis, surveillance, autonomous driving, systems maintenance, cyber-security, finance, quality assurance etc. which are in dire need of automation to help them function reliably since there is no room for error and automation in these areas can aid in reducing the time and effort spent on monotonous human tasks and to put them to better use. Conventional deep learning methods involve a lot of time and effort in terms of hyper-parameter tuning to find the best solution and this process often involves multiple iterations of development for meagre improvements, restricting generic machine learning frameworks to be put into application right-away. The modification we propose help readily improve the performance of Convolutional AutoEncoders (CAE) and potentially other CNN architectures for automated anomaly detection where the model can be directly deployed for online learning accompanying better learning capabilities with minimal human intervention.

Most anomaly detection systems are semi-supervised or unsupervised because of the evident scarcity in labelled anomalous data. Deep learning has emerged as a predominant solution for computer vision problems owing to the ability to perceive complex data and their super-human performance. In this work, we focus on the reconstruction error based method of AD using CAEs that learn the common attributes of variation from normal data to reconstruct them perfectly. Model complexity increases the computation cost and leads to the reconstruction of abnormal samples which can deteriorate the performance. An ideal AD algorithm should inherently learn the discriminating patterns between the normal and anomalous samples. It is possible to achieve optimal performance using vanilla AEs but a huge effort is needed to be spent in search of the right hyper-parameters. To alleviate this burden of searching for the optimal architectures that result in small performance gains, we propose a novel improvement that can potentially be incorporated into any convolutional architectures without any computational burden. We propose a *convolution-based softmax input soft-attention mechanism* which is capable of learning the constitution of the input images thereby helping the model focus on the essential features for optimal representations. Moreover, its differentiable nature enables training through back-propagation. The proposed input attention helps in improving the learning capabilities of CAEs and producing consistent performance improvements with minimal effort.

2. RELATED WORK

Deep Learning is the most adopted solution for visual anomaly detection due to its various advantages [2] and AEs have accomplished state-of-the-art (SOTA) performance on various anomaly detection tasks. Reconstruction-based methods often employ a variant of AE [3, 4] architecture to learn the notion of normality from normal data samples and their inability to reconstruct abnormal data is utilized. A major problem that reconstruction-based CAEs suffer from is the occasional reconstruction of anomalies that hinder their performance. [5] presents a method to limit the reconstruction capacity of AEs by introducing negative samples for discrim-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

inative learning. But there is a need for a method of better discriminative learning without need for data segregation to enable complete automation in applications. This is where attention mechanisms could play a vital role.

Attention [6], inspired by the ability of humans to focus on important parts of complex scenarios to make insights out of them, was first introduced to help neural networks concentrate on the vital portions of the data and the idea has been widely adopted for CNNs too. For visual attention, [7] introduced the concept of visual glimpse using which portions of an image were sequentially analysed region-wise for visual tasks. Later, [8] came up with object localization using reinforcement learning to detect objects region-wise and then, [9] applied soft-attention for action recognition in images followed by [10] for object localization. [11] introduced SCA-NN which uses both spatial and channel-wise attention for image captioning. All the above works acted as precursors of introducing attention into CNNs. In [12], the authors propose a modification to VGG with multiple attention blocks producing attention maps from which compatibility scores of global representations are calculated and aggregated in a fully connected dense layer for image classification. Then, [13] devised a self-attention mechanism and reformulated convolution operations by considering the relationship of pixels with their neighbours based on covariances. Later, Squeeze Excitation networks [14] were introduced for object detection achieving SOTA results. Recently, Attention Augmented Convolutions [15] was introduced as a replacement to normal convolutions where they use a relative self-attention mechanism for images. Squeeze excitation and attention augmented convolutions will be extensively dealt with, in the later sections. Closely related to our work, [16] put forward MAMA-net which is an AE network for image anomaly detection which contains a multi-scale sampler and hash-coding mechanism to retrieve similar encodings using hamming distance. Most of the above methods involve complex modifications that increase the computational complexity of the model whereas our primary focus is on developing simpler, modular improvements to readily increase the performance of any convolutional AE model.

3. METHODS

In this section, we describe the methodologies and experiments related to this work. We define a *Base-CAE* that accepts raw images (without special pre-processing) of shape 128x128 normalized between 0 and 1 with 5 convolutional and 5 transpose convolutional layers each followed by batch normalization and ReLU activation. All the layers consist of strided convolution operations with 3x3 kernels. We found that strided convolutions were better than pooling and upsampling layers as they facilitate learning. The number of kernels in each layer is 64,64,64,96,96 and the embedding dimension is 1536x1. The decoder is a mirrored-replica of the encoder with padding to adjust the output shape. Such a simple model is chosen to show the effectiveness of our approach and that performance comparable to SOTA can be achieved.

In CNNs, input propagation through multiple convolutional layers leads to learning by kernels at a local neighbourhood level and the spatial information learnt is abstracted into scalar values in subsequent layers leading to the loss of a global context. Larger kernels and deeper networks can improve the abstraction process at a high computational cost. Hence, attention can come in handy by guiding CNNs to focus on sections of the inputs that are crucial in making decisions relating to the context in the input thereby improving the performance readily without much increase in computational complexity. Attention in computer vision operates by augmenting the important parts of the image while attenuating the other parts, to emphasize the relative importance of the essential input features over the others. We focus on soft-attention mechanisms that fade parts of inputs without completely discarding the rest. The emphasis of effective attention mechanisms is on differentiability, to help learning at local and global levels with good approximation capabilities for efficient abstraction.

3.1. Attention Augmented Convolutional AutoEncoder

Attention Augmented Convolution (AAC), introduced in [15] is a convolutional self-attention mechanism that operates in a global context used primarily to augment learning for discriminative visual tasks and hence can be incorporated into CAEs. The input images are fed to a convolutional layer and a multi-head self-attention layer simultaneously and the final outputs are produced by concatenating their respective outputs while maintaining translation equivariance to retain the positional information in images [15]. Two important parameters determine the overall performance - d_k, d_v which are key depth and attention channels respectively. We replace all the convolutional layers in the encoder of the Base-CAE with AAC layers and it results in parameter increase by *17.7%*. We refer to this model as *AA CAE*.

3.2. Squeeze Excitation Convolutional AutoEncoders

Squeeze Excitation (SE) Network [14] introduces channel-wise attention to learn channel inter-dependencies by adaptive weighing of the feature map in each channel that helps the network to analyse the importance of each feature map at any layer. The 'squeeze' part of the network compresses each feature map into a scalar value and the 'excitation' part of the network is made of two fully-connected layers to process the squeezed vector into a *Sigmoid* activated vector denoting the importance of each channel and channels are scaled accordingly. For our experiment, SE blocks are added after each convolution and transpose convolution layers and the parameter increase is by *0.85%*. We refer to this model as *SE CAE*.

3.3. Proposed CAE with learnable input soft-attention

To alleviate accidental reconstruction of anomalies leading to performance deterioration, CAEs should learn discriminating features at a global scale so that the latent embeddings help only to reconstruct normal parts of the image precisely. As attention can inherently help the network focus on important features and since convolutional layers have appealing properties like universal approximating capabilities and learnable parameters, we use convolutional layers for the input soft-attention mechanism we propose. It consists of two convolutional layers with different kernel sizes k_1, k_2 where, typically $k_1 < k_2$ so that k_1 captures elementary attributes in the local neighborhood, projected over multiple channels C_p in f_{conv1} and k_2 at a global scale and compresses it to the original number of channels in f_{conv2}.

We create two variants of the mechanism - one trainable through a modified loss function and another self-contained softmax attention mechanism. The former convolution attention CAE, which we refer to as *CA CAE* consists of the convolutional layers followed by batch normalization and *Sigmoid* activation to produce the output as probability scores of importance for each pixel as an adaptive weighting mechanism to produce the final *attention map* x_{CAM}. To achieve the training objective of augmenting essential parts of the image while attenuating the rest, we add a constraint to reduce the norm of the weights in the final layer by reducing the norm of the attention map of the mechanism in the loss function with a penalty factor λ as in equation 1. We found that using $k_1 = 3, k_2 = 5, C_p = 64, \lambda = 1 \times 10^{-6}$ were effective on all the datasets in our experiments. The performance saturates with an increase in C_p and λ should be chosen such that the model doesn't lose vital information due to heavy penalty leading to diminished weights.

$$
\begin{aligned}
&x_{CAM} = Sigmoid(x_{BN}) \\
&L_{CAM} = L_{MSE} + \lambda||x_{CAM}|| \\
&S_w, S_h = AxisWiseSummation(x_{BN}) \\
&x_{SAM} = \sigma(S_w)\sigma(S_h) \\
&\hat{x} = x \odot x_a \text{ where } x_a \epsilon \{x_{CAM}, x_{SAM}\}
\end{aligned} \tag{1}
$$

The second softmax (σ) variant, *SCA CAE* alleviates the need for parameter tuning with *an important modification yielding better results*. The components till the BatchNorm layer are retained and its output x_{BN} is summed width-wise and height-wise to produce two vectors of size *1*x*H*x*C* and *W*x*1*x*C* respectively. These vectors represent the overall context value of the pixels width-wise and height-wise and applying *softmax* (σ) on them results in a probability distribution of importance along their respective axes. We then multiply the two tensors to get *W*x*H*x*C* again and this *softmax attention map*, x_{SAM} results in bands of probabilities for neighbourhoods with important features in the input image. Scaling/normalizing x_{SAM} can help in better activation visualization at intermediate layers but using x_{SAM} as such helps in better performance although the activation values are very low. The attention map $x_a \epsilon \{x_{CAM}, x_{SAM}\}$ is weighted with the input x using element-wise/ Hadamard product and sent to the CAE as inputs. We add one of the attention blocks before the input layer of our Base-CAE while retaining the rest of the architecture as such.

4. EXPERIMENTS AND RESULTS

To show the universal nature of our proposed modification, we evaluate our models on both the tasks of image and frame-level video anomaly detection with comparison to other relevant works with CAEs. We purposefully avoid hyper-parameter optimization and have chosen a simple CAE architecture to emphasize the efficacy of our proposed solution is readily improving performance. We benchmark results of Base-CAE, AA CAE, SE CAE, CA CAE and SCA CAE on 3 video datasets and 1 image dataset as in Table 1[1] and the average result of 3 runs are reported. We subject all the models to the same train sets for 300 epochs at an initial learning rate (LR) of 1 x 10^{-3} with Adam optimizer and LR decay on a plateau to aid convergence and use MSE as a loss function and as a measure of detecting anomalies. [2] We use two metrics of benchmarking AD results - area under the receiver-operator characteristics curve (AUC-ROC) which is robust and non-parametric metric to measure the separability independent of the threshold and Equal Error Rate (EER), a point on ROC where the probability of miss-classifying positive and negative samples are equal. It is apparent from Table 2 showing the performance of all models, that our proposed SCA CAE consistently outperforms other models in all our experiments, especially the baseline model Base CAE. Figure 1 shows the inputs and reconstructions of different models on a variety of datasets that are required for the analysis.

On *UCSD1 and UCSD2*, SCA CAE reconstructed anomalies like a cart into two persons as seen in Figure 1 and was able to detect cycles as anomalies whereas the other models

[1] train and test sets were used as such from the respective original sources for fair benchmarking with other works

[2] Our PyTorch codebase containing all the experiments is available at `https://github.com/ambareeshravi/Attention-AD-AE`.

Dataset	Type	Train	Test
HAM10000 [17]	Skin Lesion images	6600 images	3415 images
CUHK Avenue [18]	Surveillance videos	16 videos	21 videos
UCSD Ped 1,2 [19]	Pedestrian videos	34,36 videos	16,12 videos
Subway [20] Entrance, Exit	Surveillance videos	15,5 minutes	5,40 minutes

Table 1: Details of the anomaly detection datasets

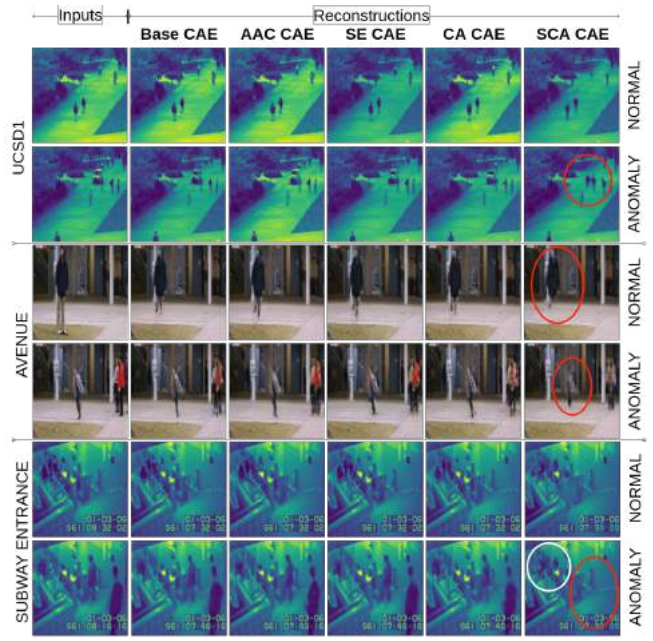


Fig. 1: Visualizing reconstructions of normal and anomalies [*Cart, kid jumping, static person*] in comparison to the inputs on *UCSD1, Avenue and Subway Entrance* respectively

were able to reconstruct cycle almost perfectly showing their inability to distinguish the anomalous object using their learnt knowledge. The overall improvements in AUC-ROC score were 3% and 5% on UCSD 1 and 2 respectively. On *Avenue*, SCA CAE was able to remove a part of people when they appeared too close to the camera as shown in Figure 1. Moreover, SCA CAE was able to identify static anomalies like a bag in the bottom-left corner of the frame correctly while the other models failed and achieving an improvement of 8% AUC-ROC. On *Subway Entrance and Exit* datasets, an intriguing fact that we observed is that the SCA CAE completely removed a person too close to the camera while preserving the remaining people seated in the background and a person towards the turnstile as shown in Figure 1 while other models completely removed people in the frame. This shows the ability of the SCA CAE to learn normality well under context and distinguish the abnormalities using the learnt knowledge. *HAM10000* is inherently complicated and tangible reasoning without domain knowledge is hard and hence we rely solely on the performance metrics for comparison as the visual analysis is difficult. We compare the performance of our models to other SOTA works such as [21] which uses CAE and [22] Variational CAEs with class-wise mean AUC-ROC, though their architectures are slightly more complex than ours. Though the reconstructions are slightly blurry, SCA CAE can identify anomalies better than the other models due to the reduction in the overall reconstruction capability of the CAE alleviating the problem of partially or fully reconstructing anomalies and it is highly sensitive to the variations inputs and in a few cases, it was even able to detect

Dataset	Model	AUC-ROC % ↑	EER % ↓
AVENUE	Base-CAE	81.60	25.97
	AA CAE	78.08	27.84
	SE CAE	80.52	24.86
	CA CAE	79.75	25.93
	SCA CAE	**89.67**	**15.76**
	CAE [21]	70.2	25.1
SUBWAY ENTR.	Base-CAE	74.28	33.38
	AA CAE	74.61	33.28
	SE CAE	72.08	32.13
	CA CAE	73.85	33.10
	SCA CAE	*74.88*	*31.41*
	CAE [21]	**94.3**	**26.0**
SUBWAY EXIT	Base-CAE	95.68	11.11
	AA CAE	95.72	11.10
	SE CAE	95.53	11.11
	CA CAE	95.60	11.11
	SCA CAE	**97.75**	**5.92**
	CAE [21]	80.7	9.9
UCSD1	Base-CAE	68.32	37.59
	AA CAE	66.06	38.54
	SE CAE	66.49	38.45
	CA CAE	67.54	38.00
	SCA CAE	*71.09*	*33.28*
	CAE [21]	**81.0**	**27.9**
UCSD2	Base-CAE	83.85	26.52
	AA CAE	83.28	26.38
	SE CAE	82.71	27.18
	CA CAE	86.89	19.36
	SCA CAE	*88.06*	*18.78*
	CAE [21]	**90.0**	**21.7**
HAM 10000	Base-CAE	68.60	35.50
	AA CAE	68.78	35.19
	SE CAE	69.36	34.74
	CA CAE	69.10	35.08
	SCA CAE	*70.15* ***[Mean 76.69]***	*34.45*
	VAE [22]	**Mean 77.9**	**N/A**

Table 2: Evaluation results of our experiments on datasets

people walking in the opposite direction which is an anomaly category in UCSD and Subway datasets though the temporal aspect was not taken into account while training.

5. CONCLUSION

In this paper, we have studied the potential application of several attention mechanisms to CAE architectures and their ability to augment the anomaly detection performance which can help in speeding up deployment for automated online, self-learning applications with little human intervention. We have also proposed a novel input attention mechanism to readily improve the performance in CAEs. We have also discussed in detail, the nuances and reasoning behind the working of the mechanisms with comprehensive experimental analysis on multiple benchmark datasets demonstrating the superiority of the proposed mechanism on image and video anomaly detection tasks. For future work, we are interested in analysing the performance of our proposed mechanisms on other visual tasks and also in prediction based Anomaly Detection tasks.

6. REFERENCES

[1] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58, 2009.

[2] Raghavendra Chalapathy and Sanjay Chawla. Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*, 2019.

[3] Yoshua Bengio, Pascal Lamblin, Dan Popovici, Hugo Larochelle, et al. Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, 19:153, 2007.

[4] Manassés Ribeiro, André Eugênio Lazzaretti, and Heitor Silvério Lopes. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recognition Letters*, 105:13–22, 2018.

[5] Asim Munawar, Phongtharin Vinayavekhin, and Giovanni De Magistris. Limiting the reconstruction capability of generative neural network using negative learning. In *2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE, 2017.

[6] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.

[7] Volodymyr Mnih, Nicolas Heess, Alex Graves, and Koray Kavukcuoglu. Recurrent models of visual attention. *arXiv preprint arXiv:1406.6247*, 2014.

[8] Juan C Caicedo and Svetlana Lazebnik. Active object localization with deep reinforcement learning. In *Proceedings of the IEEE international conference on computer vision*, pages 2488–2496, 2015.

[9] Shikhar Sharma, Ryan Kiros, and Ruslan Salakhutdinov. Action recognition using visual attention. *arXiv preprint arXiv:1511.04119*, 2015.

[10] Eu Wern Teh, Mrigank Rochan, and Yang Wang. Attention networks for weakly supervised object localization. In *BMVC*, pages 1–11, 2016.

[11] Long Chen, Hanwang Zhang, Jun Xiao, Liqiang Nie, Jian Shao, Wei Liu, and Tat-Seng Chua. Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5659–5667, 2017.

[12] Saumya Jetley, Nicholas A Lord, Namhoon Lee, and Philip HS Torr. Learn to pay attention. *arXiv preprint arXiv:1804.02391*, 2018.

[13] Xiaolong Wang, Ross Girshick, Abhinav Gupta, and Kaiming He. Non-local neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7794–7803, 2018.

[14] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[15] Irwan Bello, Barret Zoph, Ashish Vaswani, Jonathon Shlens, and Quoc V Le. Attention augmented convolutional networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3286–3295, 2019.

[16] Yurong Chen, Hui Zhang, Yaonan Wang, Yimin Yang, Xianen Zhou, and QM Jonathan Wu. Mama net: Multi-scale attention memory autoencoder network for anomaly detection. *IEEE Transactions on Medical Imaging*, 2020.

[17] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1):1–9, 2018.

[18] Cewu Lu, Jianping Shi, and Jiaya Jia. Abnormal event detection at 150 fps in matlab. In *Proceedings of the IEEE international conference on computer vision*, pages 2720–2727, 2013.

[19] Vijay Mahadevan, Weixin Li, Viral Bhalodia, and Nuno Vasconcelos. Anomaly detection in crowded scenes. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1975–1981. IEEE, 2010.

[20] Amit Adam, Ehud Rivlin, Ilan Shimshoni, and Daviv Reinitz. Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE transactions on pattern analysis and machine intelligence*, 30(3):555–560, 2008.

[21] Mahmudul Hasan, Jonghyun Choi, Jan Neumann, Amit K Roy-Chowdhury, and Larry S Davis. Learning temporal regularity in video sequences. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 733–742, 2016.

[22] Yuchen Lu and Peng Xu. Anomaly detection for skin disease images using variational autoencoder. *arXiv preprint arXiv:1807.01349*, 2018.

ANOMALY-AWARE FEDERATED LEARNING WITH HETEROGENEOUS DATA

Zheng Chen, Chung-Hsuan Hu, and Erik G. Larsson

Department of Electrical Engineering, Linköping University, Sweden.
{chung-hsuan.hu, zheng.chen, erik.g.larsson}@liu.se

ABSTRACT

Anomaly detection plays a critical role in ensuring the robustness and reliability of federated learning (FL) systems involving distributed implementation of stochastic gradient descent (SGD). Existing methods in the literature usually apply norm-based gradient filters in each iteration and eliminate possible outliers, which can be ineffective in a setting with heterogeneous and unbalanced training data. We propose a heuristic yet novel scheme for adjusting the weights in the gradient aggregation step that accounts for two anomaly metrics, namely the relative distance and the convergence measure. Simulation results show that our proposed scheme brings notable performance gain compared to norm-based policies when the agents have distinct data distributions.

Index Terms— Federated learning, anomaly detection, gradient aggregation rule, fault tolerance

1. INTRODUCTION

Federated learning (FL) is a newly emerged concept for collaborative training among multiple agents which participate in a global machine learning (ML) task without sharing their privacy-sensitive data to a centralized server [1–3]. In a server-based FL system, a server is responsible for aggregating and updating the model parameters in an iterative process. In every global iteration, each agent trains the current learning model based on its raw data and sends the model parameter updates (e.g., gradient information) to the parameter server. The server aggregates the received updates using some weighted averaging policy and obtains an updated parameter vector, which will then be distributed to the agents in the next iteration. Since the server does not have direct access to the training data, the outcome of the training process can be severely affected by abnormal agent behaviors including intended attacks and unintended machine failures [4–6]. For instance, it is shown in [7] that a single attacker can prevent convergence of a distributed stochastic gradient descent (SGD) system if a linear combination rule is applied for gradient aggregation.

This work was supported in part by Centrum för Industriell Informationsteknologi (CENIIT), Excellence Center at Linköping - Lund in Information Technology (ELLIIT), and Knut and Alice Wallenberg (KAW) Foundation.

Many existing works have studied the robustness aspect of distributed SGD systems in terms of Byzantine Fault Tolerance (BFT) [8]. A distributed system is said to achieve BFT if the non-faulty clients can still converge to an agreement in the presence of some Byzantine attackers that intend to disturb the system from converging. Under the assumption of Independently and Identically Distributed (IID) data at different agents, several gradient filtering schemes have been proposed, such as *Krum* [7], *Median* [9], and *comparative gradient clipping (CGC)* [10]. The core idea behind these schemes is to eliminate the outliers based on the gradient norms or the distance between the gradient vectors. Using the concept of consensus optimization, a regularization-based technique is proposed in [11] for achieving Byzantine-robust distributed learning with heterogeneous datasets. All the aforementioned methods consider only the gradient or model updates within each iteration, and the evolution of the model updates over different iterations has never been investigated for anomaly detection. Nevertheless, for an FL system, ensuring the system convergence is not sufficient for ensuring the learning accuracy, as the ML model might converge to a sub-optimal solution. In a realistic scenario with non-IID and unbalanced training data, an efficient anomaly detection scheme should take into account the variation among the model updates from different agents caused by the data heterogeneity.

In this work, we propose a heuristic anomaly-aware update aggregation scheme for an FL system with heterogeneous training data distribution. In the proposed scheme, the server adjusts the weight coefficients for the received model updates from different agents, based on the relative distances between the gradient vectors and the temporal evolution of the gradient update from each agent.

2. SYSTEM MODEL

We consider an FL system with one server and a set of N distributed agents, where all the agents are connected to the server to participate in the same learning task. Each agent $i \in \{1, \ldots, N\}$ has its own local data set $\xi_i = \{(\mathbf{x}, \mathbf{y})_{i,1}, \ldots, (\mathbf{x}, \mathbf{y})_{i,D_i}\}$, where $\mathbf{x} \in \mathbb{R}^{d_x}$ is the input data with dimension d_x, $\mathbf{y} \in \mathbb{R}^{d_y}$ is the output data with dimension d_y, and $D_i = |\xi_i|$ is the size of local data set. The size of the entire set of training data is $D = \sum_{i=1}^{N} D_i$. The global

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

learning model is parameterized by the vector $\boldsymbol{\theta} \in \mathbb{R}^d$, where d is the dimension of the parameter vector.

The objective of the learning task is to find the parameter vector that minimizes the empirical loss function over the entire data set. The global loss function is defined by

$$F(\boldsymbol{\theta}) = \frac{1}{D} \sum_{i=1}^{N} \sum_{j=1}^{D_i} f_{i,j}(\boldsymbol{\theta}), \tag{1}$$

where $f_{i,j}(\boldsymbol{\theta})$ is the sample-wise loss function of the j-th data sample at the i-th agent. Define the local loss function at each agent i as

$$f_i(\boldsymbol{\theta}) = \frac{1}{D_i} \sum_{j=1}^{D_i} f_{i,j}(\boldsymbol{\theta}). \tag{2}$$

The global loss function can be written as the weighted average of the local loss functions, i.e.,

$$F(\boldsymbol{\theta}) = \sum_{i=1}^{N} \frac{D_i}{D} f_i(\boldsymbol{\theta}). \tag{3}$$

2.1. Introduction to Federated Averaging

According to the original version of Federated Averaging (FedAvg) algorithm proposed in [2], a typical FL process is an iterative collaborative learning process that consist of two main parts: local training and global aggregation. At the beginning of t-th global iteration, the server distributes the current model parameter vector $\boldsymbol{\theta}(t)$ to the participating agents. After receiving the parameter vector $\boldsymbol{\theta}_i(t) = \boldsymbol{\theta}(t)$, agent i performs one or several steps of SGD on its local data set and obtains an updated parameter vector $\boldsymbol{\theta}_i(t+1)$. The difference between the parameter vector before and after its local iterations, referred to as the model update, is defined as

$$\mathbf{u}_i(t) = \boldsymbol{\theta}_i(t+1) - \boldsymbol{\theta}_i(t). \tag{4}$$

If we consider that the local training process consists of only one step of SGD based on a randomly selected mini-batch, the model update information corresponds to

$$\mathbf{u}_i(t) = -\eta_i \nabla f_i(\boldsymbol{\theta}_i(t)), \tag{5}$$

where $\nabla f_i(\boldsymbol{\theta}_i(t))$ is the gradient of the local loss function for the current parameter vector $\boldsymbol{\theta}_i(t)$ calculated over a randomly selected mini-batch within the local data set ξ_i. After completing the local SGD step, each agent sends back the model update (gradient information) to the server.

After receiving all the local updates, the server aggregates and updates the parameter vector by applying a weighted sum of the received model updates. The updated global parameter vector for the next iteration becomes

$$\boldsymbol{\theta}(t+1) = \boldsymbol{\theta}(t) + \sum_{i=1}^{N} w_i(t) \mathbf{u}_i(t), \tag{6}$$

where $w_i(t) = \frac{D_i}{D}$ is the weight associated with the model update from agent i.

2.2. Convergence Conditions

We define the global optimal solution as

$$\boldsymbol{\theta}^* = \arg\min F(\boldsymbol{\theta}). \tag{7}$$

If all the agents are trustworthy and accurate, the convergence of FedAvg algorithm to the optimal solution is guaranteed if the following assumptions hold.

- The local loss functions $f_i(\boldsymbol{\theta})$ for all $i \in \{1, \ldots, N\}$ are strongly convex and L-smooth.
- The local gradients $\nabla f_i(\boldsymbol{\theta}_i(t))$ for all $i \in \{1, \ldots, N\}$ have bounded variance and bounded squared norm.

Since $\nabla F(\boldsymbol{\theta}^*) = 0$, which implies that when $\boldsymbol{\theta}(t)$ approaches $\boldsymbol{\theta}^*$, we have $\sum_{i=1}^{N} w_i \mathbf{u}_i(t) \to 0$ with $\mathbf{u}_i(t) \to \mathbf{u}_i^*$. This means that the model update at each agent i converges to $\mathbf{u}_i^*$ if the convergence conditions are satisfied.

3. ANOMALY DETECTION WITH HETEROGENEOUS DATA

When the agents have distinct data distributions, it is likely that some of the model updates (stochastic gradients) have very large values even after system convergence. Conventional norm-based gradient elimination schemes for distributed SGD systems might be efficient at achieving system convergence, but the accuracy of the trained ML model will be affected. Therefore, a reasonable anomaly detection scheme for FL systems should identify suspicious agents without raising alarms for the ones that have relatively different data distribution than the others.

In this work, we consider an agent to be normal/regular if in every round the agent uploads its true gradient update to the server. Abnormal agents refer to the ones that send incorrect gradient updates, either because of machine failures or malicious attacks.

Our proposed scheme combines two anomaly metrics into the design of the weight factors in (6) for model averaging, namely the relative distance measure and the local convergence measure. Most importantly, we exploit the evolution of the gradient updates over several consecutive iterations as the most important abnormal behavior indicator.

3.1. Relative Distance Measure

For each agent i, we measure the distance between its local gradient vector and the aggregated gradient vector from the others. The distance measure of agent i in the t-th global iteration is defined by

$$d_i(t) = \left\| \mathbf{u}_i(t) - \frac{1}{N-1} \sum_{k=1, k \neq i}^{N} \mathbf{u}_k(t) \right\|. \tag{8}$$

This metric indicates how far the gradient of one agent is from the average gradient of the other agents. However, it is not sufficient to raise the alarm when one agent has a large $d_i(t)$, as it might indicate that agent i has different data distribution than the others.

3.2. Local Convergence Measure

To measure the convergence of the local gradient updates from each agent, we propose a metric based on the relative differences between the local gradients within three consecutive iterations. The local convergence measure of agent i in the t-th global iteration is defined by

$$c_i(t) = \frac{\|\mathbf{u}_i(t) - \mathbf{u}_i(t-1)\|}{\|\mathbf{u}_i(t-1) - \mathbf{u}_i(t-2)\|}. \tag{9}$$

The intuition behind this definition is that, if the local gradient from agent i shows convergence, i.e., $\mathbf{u}_i(t) \to \mathbf{u}_i^*$ when $t \to \infty$, then the local convergence measure will be strictly smaller than 1.

3.3. Update Aggregation with Anomaly Score

Similar to [12], we assign an anomaly score to every agent to quantify its level of abnormality. We define $A_i(t)$ as the anomaly score of agent i in the t-th global iteration, given by

$$A_i(t) = c_i(t) \cdot \frac{d_i(t)}{\max\limits_{j=1,\ldots,K}\{d_j(t)\}}, \tag{10}$$

The anomaly score has the following properties:

- $A_i(t) \geq 0$.
- $A_i(t)$ is positively correlated to both $c_i(t)$ and $d_i(t)$.
- When $c_i(t) \simeq c_i(t)$ for all $i, j \in \{1, \ldots, N\}$, $A_i(t)$ is mostly limited by the relative distance measure.

Based on the definition above, we propose an update aggregation rule where the weight coefficient for each agent is given by the following metric

$$w_i(t) = \frac{D_i \exp(-A_i(t))}{\sum_{j=1}^{N} D_j \exp(-A_j(t))}, \tag{11}$$

such that $\sum_{i=1}^{N} w_i(t) = 1$. The weight coefficient $w_i(t)$ is a monotonically decreasing function of $A_i(t)$, which implies that an agent with a higher anomaly score will have a lower weight in the update aggregation process.

4. SIMULATION RESULTS

The proposed method is implemented and tested with a nonlinear polynomial function fitting problem. We consider one server and $N = 10$ agents in the system, where each agent $i \in \{1, ..., N\}$ holds a local training data set ξ_i with sample size $D_i = |\xi_i| = 100$. Every data sample (x, y) with $x \in \mathbb{R}$ and $y \in \mathbb{R}$ is generated by a polynomial function $y = g(x, \boldsymbol{\theta}) + w$ parameterized by $\boldsymbol{\theta} \in \mathbb{R}^3$, with $w \sim N(0, 1)$ denoting the random noise. In each global iteration, the agents perform one step of SGD and return the model update information as defined in (5). The empirical loss function is defined by the mean square error (MSE) of the trained model. For the j-th data sample at the i-th agent, the sample-wise loss function is defined as

$$f_{i,j}(\boldsymbol{\theta}) = |g(x_{i,j}, \boldsymbol{\theta}) - y_{i,j}|^2. \tag{12}$$

We consider an extreme case of non-IID data distribution among the agents. For each agent i, the training data samples are generated within the range $x \in [x_{\min,i}, x_{\max,i}]$, where all the data ranges are non-overlapping. This implies that each agent has a unique distribution of data samples. In the simulations, we consider only one malicious agent, with two types of attack described as follows.

- **Gradient ascent**. A malicious client k computes its local gradient and flips the sign of $\mathbf{u}_k(t)$.
- **Random perturbation**. A malicious client k replaces its true gradient $\mathbf{u}_k(t)$ by a random number $n(t)$ generated from the distribution $N(0, \sigma_n^2)$, with $\sigma_n^2 = 0.25$ used in the simulations.

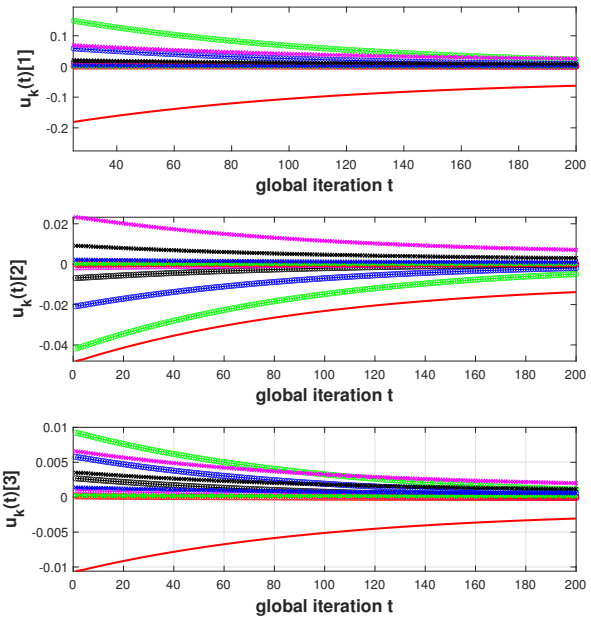


Fig. 1: The evolution of the gradient updates in the presence of one malicious agent with "gradient ascent" attack.

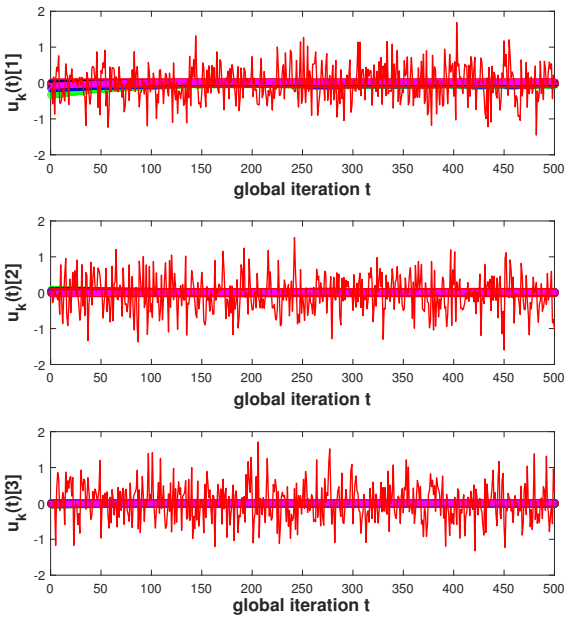


Fig. 2: The evolution of the gradient updates in the presence of one malicious agent with "random perturbation" attack.

Figs. 1 and 2 show the evolution of the local model updates $\mathbf{u}_k(t)$ from all the agents under gradient ascent and random perturbation attacks. From Fig. 1, we can see that due to the heterogeneity of training data, some of the gradient updates from the regular/normal agents can be quite large, which could be easily excluded in the aggregation process if we rely on norm-based gradient filtering methods. From Fig. 2, we see that random perturbations can be easily detected by observing the temporal evolution of the gradient updates.

In Figs. 3 and 4, we plot the MSE of our proposed method, marked as "grdtAsct:Proposed" and "grdtAwgn:Proposed", when the system is under gradient ascent and random perturbation attacks, respectively. For performance comparison, we also plot the simulation results for two reference methods including "Multi-Krum" in [7] and "Median" in [9]. Moreover, "normal: Baseline" represents the best-case performance without any malicious agent, and "No Action" represents the case where the system is under attack but no anomaly detection scheme is deployed. From these two figures, we observe that our proposed method is efficient at eliminating the effect of potential attacks and achieves superior performance compared to norm-based gradient filtering techniques. In Fig. 4, the performance gain compared to the case without anomaly detection scheme is less significant, since SGD is by default a noisy process with estimation noise introduced in every iteration. It is also worth noting that even though the reference methods have theoretically provable convergence

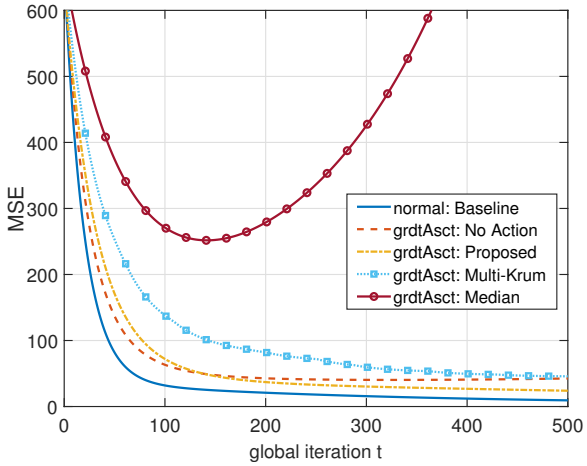


Fig. 3: Performance comparison for the scenario with "gradient ascent" attack.

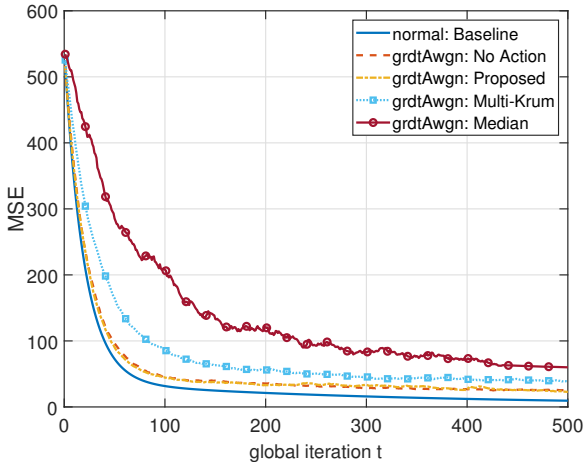


Fig. 4: Performance comparison for the scenario with "random perturbation" attack.

guarantee under the IID data assumption, their performance is less promising in a realistic FL scenario with heterogeneous data.

5. CONCLUSIONS

In this work, we proposed an anomaly-aware gradient aggregation rule for FL systems with heterogeneous training data distributions at the participating agents. Compared to commonly adopted norm-based gradient filtering methods, our proposed scheme combined the relative distance measure and the temporal evolution of the local gradient updates over consecutive iterations as two important anomaly indicators. Simulation results showed the efficiency of our proposed scheme at eliminating the effect of potential attacks while respecting the data heterogeneity of the non-faulty agents.

6. REFERENCES

[1] J. Konečnỳ, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated optimization: Distributed machine learning for on-device intelligence," *arXiv preprint arXiv:1610.02527*, 2016.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial Intelligence and Statistics*. PMLR, 2017, pp. 1273–1282.

[3] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[4] E. M. El Mhamdi, R. Guerraoui, and S. L. A. Rouault, "The hidden vulnerability of distributed learning in byzantium," in *International Conference on Machine Learning*, 2018.

[5] L. Lyu, H. Yu, X. Ma, L. Sun, J. Zhao, Q. Yang, and P. S. Yu, "Privacy and robustness in federated learning: Attacks and defenses," *arXiv preprint arXiv:2012.06337*, 2020.

[6] M. S. Jere, T. Farnan, and F. Koushanfar, "A taxonomy of attacks on federated learning," *IEEE Security Privacy*, vol. 19, no. 2, pp. 20–28, 2021.

[7] P. Blanchard, E. M. El Mhamdi, R. Guerraoui, and J. Stainer, "Machine learning with adversaries: Byzantine tolerant gradient descent," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 118–128.

[8] Y. Dong, J. Cheng, M. J. Hossain, and V. C. M. Leung, "Secure distributed on-device learning networks with byzantine adversaries," *IEEE Network*, vol. 33, no. 6, pp. 180–187, 2019.

[9] Y. Chen, L. Su, and J. Xu, "Distributed statistical machine learning in adversarial settings: Byzantine gradient descent," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 1, no. 2, pp. 1–25, 2017.

[10] N. Gupta and N. H. Vaidya, "Fault-tolerance in distributed optimization: The case of redundancy," in *Proceedings of the 39th Symposium on Principles of Distributed Computing*, 2020, pp. 365–374.

[11] L. Li, W. Xu, T. Chen, G. B. Giannakis, and Q. Ling, "RSA: Byzantine-robust stochastic aggregation methods for distributed learning from heterogeneous datasets," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 1544–1551.

[12] S. Li, Y. Cheng, Y. Liu, W. Wang, and T. Chen, "Abnormal client behavior detection in federated learning," *arXiv preprint arXiv:1910.09933*, 2019.

ONLINE UNSUPERVISED LEARNING FOR DOMAIN SHIFT IN COVID-19 CT SCAN DATASETS

Nicolas Ewen[⋆]*, Naimul Khan*[⋆]

[⋆] Electrical, Computer, and Biomedical Engineering, Ryerson University, Toronto, ON

ABSTRACT

Neural networks often require large amounts of expert annotated data to train. When changes are made in the process of medical imaging, trained networks may not perform as well, and obtaining large amounts of expert annotations for each change in the imaging process can be time consuming and expensive. Online unsupervised learning is a method that has been proposed to deal with situations where there is a domain shift in incoming data, and a lack of annotations. The aim of this study is to see whether online unsupervised learning can help COVID-19 CT scan classification models adjust to slight domain shifts, when there are no annotations available for the new data. A total of six experiments are performed using three test datasets with differing amounts of domain shift. These experiments compare the performance of the online unsupervised learning strategy to a baseline, as well as comparing how the strategy performs on different domain shifts. Code for online unsupervised learning can be found at this link: https://github.com/Mewtwo/online-unsupervised-learning

***Index Terms*—** online unsupervised learning, self supervision, CNN, transfer learning, neural network, medical image, COVID-19, CT scan

1. INTRODUCTION

In this study we aim to determine whether unsupervised online learning can increase classification performance of convolutional neural networks (CNNs) on COVID-19 CT scan datasets [1]. We will explore the scenario where the target datasets have no annotations, and have slight domain shifts from the available training data. A strategy for unsupervised online learning for COVID-19 CT scans is proposed, and its performance is evaluated on three different test sets.

CNN models have been effective for classification on medical imaging datasets [2][3][4][5][6]. CNN models can be hard to train on medical imaging datasets because they require large amounts of annotated data. Large amounts of annotated data may not be available for a variety of reasons, including cost and availability of expert annotators [5][6]. Data augmentation, transfer learning, and self supervision are methods that can be used to help increase CNN performance when there is a small amount of annotated data [5][6][7], however, unsupervised learning methods are needed when there are no annotations [8].

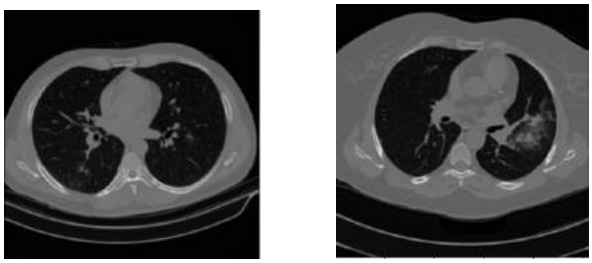


Fig. 1. Example of a CT scan with Covid-19 (left), and with CAP (right).

Performing transfer learning from a CNN pre-trained on ImageNet has improved model performance for COVID-19 CT scan datasets over training from scratch in a number of papers [9][4]. Some recent works further improved classification performance by using self supervision before transfer learning [8][10][11].

One method that is used to deal with few available annotations for training is transfer learning. Transfer learning is performed by taking a pre-trained network and then fine tuning it using the target dataset onto the target task [4]. This method allows simple and reusable features in the early layers of the network to benefit from training on an unrelated dataset. Since early layers are generally more reusable, the network will not have to train them as much, thus reducing the amount of annotated data needed to train the network [4][12].

Another method of dealing with a limited amount of available annotations for training is self supervised learning, which is a form of unsupervised learning. Self supervised learning uses unlabelled or automatically labelled data to pre-train the network to learn useful feature semantics in the target images [8]. Once the network has been pre-trained, it is then fine tuned onto the target dataset using transfer learning. This method allows the network to learn on limited labelled training data since less data will be needed after the network has already learned some useful feature semantics [8].

Online unsupervised learning is a combination of online machine learning and unsupervised learning. Online machine

learning is a type of machine learning where the model continuously updates itself with new data as the new data arrives. Unsupervised learning allows models to learn from data without any expert annotations [8]. Online unsupervised learning allows the model to continuously improve as new data without annotations comes in [13].

Online unsupervised learning is a field of machine learning that can help predictive models adapt to new situations. New illnesses and screening methods, combined with a lack of expert annotators may cause domain shifts in incoming data, with no labels. Online unsupervised learning can help train models under such circumstances, where other techniques may have to wait for more data or annotations [13].

Our main contribution in this work is to highlight and demonstrate an online unsupervised learning strategy. While the idea of using unsupervised online learning to increase classification performance is not new [14], to the best of our knowledge it has not been done for medical imaging on a COVID-19 CT scan dataset.

We felt online learning was a good approach for COVID-19 CT scans because as more data is collected, the models can be updated. We decided to model what this might look like in practise by dividing each dataset into quarters and performing the online updates after each quarter. This would allow real-time results, while also continuously increasing classification performance.

The three test datasets are good to use for this experiment. The first test dataset comes from the same settings as the training and validation set. This set should be a benchmark for how much the model improves only due to extra training and data, since there is no domain shift. This means that with the first test set, we are performing semi-supervised learning, as opposed to unsupervised learning. The second test dataset is only COVID-19, and healthy patients, but with low dosage. This should be a slight domain shift. The third test dataset has COVID-19, CAP, and healthy patients. The patients also have a heart condition, and the dosage and slice thickness vary. This test set has the largest domain shift from the training and validation sets. Combined, the results from these three test datasets should show how well our proposed strategy of online unsupervised learning adapts to slight domain shifts.

This work is very important because if a method of unsupervised online learning can be used to increase classification performance under domain shift, then new models will not have to be trained from scratch each time the domain shifts slightly. It also means that a model can be updated as new data becomes available, without having to wait for an expert to annotate the new data.

We believe that our proposed strategy can be used in practise in the real world. Hospitals performing CT scans for COVID-19 can use their models to produce real time predictions, and the models update themselves as more data comes in. Since this strategy does not use annotations on the new data coming in, the model can be updated at a rate dependant on the rate of CT scans. As a demonstration of how this strategy can be used, we divide our three test datasets into four quarters. Each quarter will be treated as patients coming in sequential order. This means that the first quarter will be evaluated with our base models. The data from the first quarter will then be used to update the models. Then the second quarter will be evaluated using the updated models, and so on. By dividing our three test datasets into quarters, we are able to run six experiments to test how well this method of online unsupervised learning adapts to slight domain shifts in COVID-19 CT scan datasets.

2. METHODOLOGY

2.1. Dataset

The dataset we used for this paper was the dataset used in the SPGC COVID-19 competition [1]. This dataset is a dataset of chest CT scan images organized by patient. The patients can be in one of three classes: Healthy, COVID-19, or CAP. There are a total of 307 patients in the training and validation sets. There are 76 healthy patients, 171 COVID-19 patients, and 60 CAP patients. Patient-level labels are provided by three radiologists, who have greater than 90 percent agreement. Images were taken under different circumstances, including different medical centres, scanners, using different slice thicknesses, effective mA, and exposure time. 55 of the COVID-19 patients, and 25 of the CAP patients have slice-level labels provided by a single radiologist. There are about 5000 slices labelled positive for infection, and about 18,500 labelled negative for infection.

In addition to the training and validation sets provided, there are also three test sets. The first test set is from patients classified as healthy, COVID-19 positive, and CAP positive, and comes from the same distribution as the training and validation sets. The second test set from patients classified as healthy and COVID-19 positive only, and a lower dosage was used for the scans. The third test set comes from patients classified as healthy, COVID-19 positive, and CAP positive. The scans were administered under various settings. The healthy patients in this test set also had an unrelated disease.

2.2. Pre-processing

The SPGC COVID-19 dataset has pre-set training, validation, and testing sets. We used these sets and did not change them. We extracted the slice-level labels and images where slice-level annotations were provided. We kept the images and labels that were positive for COVID-19 and CAP, but we did not keep the negative ones. Instead we extracted slices from healthy patients with large lung area, in a similar manner to Rahimzadeh et al. [15]. We made an image selection algorithm to filter out images without lungs, or with small sections of lung, as well as images with lungs where much of the lung is not visible. This was done by setting an inner area of the

image, and counting darker pixels in said area. A threshold was calculated for each patient on the fly using the average number of dark pixels, and images with less dark pixels in the area than the threshold were removed. These were used for normal slices. Images were resized to 224 x 224, and rescaled to between 0 and 1. The validation set was used to tune hyperparameters. For each patient, we also saved an array of the images with large lung area.

2.3. Slice-level models

We trained two different slice-level models. The first model was trained to classify slices as healthy or not healthy. This model was trained with the healthy slices we extracted from healthy patients, and the labelled slices provided for COVID-19 and CAP patients. The second model was trained to classify unhealthy slices as either COVID-19 or CAP. This model was trained with the labelled slices provided for COVID-19 and CAP patients.

For both models, we used the same network architecture and training strategy. The only difference was the data used. The network uses a DenseNet169 base, with a dense layer with 8 nodes followed by a softmax output. Batch normalization, regularization, and dropout were used as well. Our training strategy was a two-step process. For each model, we first performed targeted self supervision in a similar manner to Ewen and Khan [16]. We made horizontally flipped copies of our training images, and trained the network to determine whether an image was flipped or not. The second step in training was to then transfer onto our target dataset.

2.4. Patient-level models

At the patient level, we used the slices for each patient that we had previously extracted with larger lung area. The chosen slices were first sent through our slice-level model that classifies the slices as healthy or unhealthy. We took the average score of the patient's softmax scores to get two average scores: a healthy score, and an unhealthy score. If the healthy score was greater than five times the unhealthy score, the patient was classified as healthy. Otherwise, the patient was classified as unhealthy. This threshold of five times was chosen after testing on the validation set.

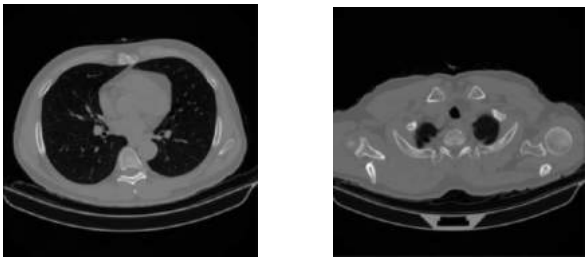


Fig. 2. Example of a healthy patient's slices with large lung area (left) and without (right).

If the patient was classified as unhealthy, then the patients' slices were sent to the next slice-level model. The chosen slices were then classified as either COVID-19 or CAP. We again took the average softmax scores to get two average scores: a COVID-19 score, and a CAP score. If the COVID-19 score was greater than the CAP score, the patient was classified as having COVID-19. Otherwise, the patient was classified as having CAP.

2.5. Adjusting patient-level thresholds

On a number of runs, the slice level models produced a large difference in the recall of the classes. For example, due to many more images with COVID-19 than CAP, the slice level model could have a recall of about 0.99 for COVID-19 images, but only 0.72 for CAP images on the validation set. This suggests that the slice level model is more likely to classify CAP incorrectly than COVID-19. This could cause incorrect classification in borderline cases.

For example, if the slice level model classified 31 slices as COVID-19, and 30 as CAP, then the patient level model will more likely classify the patient as having COVID-19. However, since the patient only has one of either CAP or COVID-19, about 30 slices have been misclassified. Given the difference in the recalls of the slice level model, it is more likely that CAP images were classified incorrectly, and that it would be better to classify this patient as having CAP.

To deal with this problem, we came up with a method of adjusting the set threshold levels in the patient level models on the fly with a multiplier. To calculate these multipliers, we took the ratio of the two recalls. In our example this is CAP recall divided by COVID-19 recall. With a CAP recall of 0.72 and a COVID-19 recall of 0.99, this would be 0.72/0.99 = 0.725. We then multiply the number of COVID-19 slices by the multiplier. This would give us 31*0.725 = 22.5 effective slices. Since this number is lower than the 30 CAP slices, the patient is then more likely to be classified as having CAP. Similarly, a second threshold multiplier was also calculated for the healthy threshold.

If performances of the respective recalls are reversed, the threshold multiplier will be greater than 1. This means that the threshold adjustment will always be made in favour of the class with lower recall, and that as the difference in recall grows, the threshold multiplier, and therefore adjustment, gets larger.

2.6. Online Unsupervised learning

We tested this method using three different COVID-19 CT scan test sets. Our "baseline" contains two networks trained on the training and validation sets, first one providing healthy/unhealthy binary slice level classification, second one

Table 1. Experiments

	Proposed Experiments	
	Test Set	***Model***
Exp. 1	Test Set 1	Baseline
Exp. 2	Test Set 2	Baseline
Exp. 3	Test Set 3	Baseline
Exp. 4	Test Set 1	Online Unsupervised
Exp. 5	Test Set 2	Online Unsupervised
Exp. 6	Test Set 3	Online Unsupervised

providing COVID/CAP slice classification on the unhealthy slices from the first model. For each test dataset, we retrained the slice level models using our online unsupervised learning scheme. This gave us 4 total models, the baseline model, and a model specifically fine tuned to its respective test dataset.

To perform the online unsupervised updates to the models, we first ran the next quarter of test data through our models to obtain predictions. From these predictions we took the confident slices. Confident slices were those with a softmax score of at least 0.9 in agreement with the patient's classification. We then used these slices, along with our label that we assigned to it during predictions, and our original training data, to retrain the model. The aim was for this to allow the model to adjust to slight domain shifts, such as lower dosage.

We used a strategy that is adjusted from what was proposed by Cao and He [14]. After predictions were generated for a batch, a new slice-level model was trained for both healthy vs unhealthy classification as well as COVID-19 vs CAP classification. These two new models were initiated from the point that the self supervision step had finished for the base models. The confident images from the test batch, along with their predictions, were added to the original training and validation sets, and then trained in the same way as the base slice-level models. This means that after receiving every test batch, two new slice-level models were trained.

2.7. Experiments

We ran six experiments in total. For each test set, we ran an experiment using the baseline method to predict the class of the patients, and another experiment using the online unsupervised method to predict the patients' class. There are three test sets, so this resulted in a total of six experiments.

3. RESULTS

The results of the experiments can be seen in Table 2. On the first test set, the baseline method got 90 percent accuracy, while the online unsupervised method got 86.7 percent accuracy.This result was unexpected, since test set 1 comes from the same distribution as the training and validation set, and the baseline performed well. A possible explanation for this result is that the test set was too small. 30 patients may not be sufficient to demonstrate the proposed method. The initial guesses from the online method are the same as for the baseline, since the models have not yet updated. This means that model performance decreased even though most patients were correctly classified initially.

On the second test set, the baseline method got 66.7 percent accuracy, and the online unsupervised method got 76.7 percent accuracy. This dataset had a small domain shift from the training and validation sets, and seemed ideal for our method. The increase in performance is promising.

The third test set had the larger domain shift from the training and validation sets. The baseline method got 63.3 percent accuracy, while the online unsupervised method got 53.3 percent accuracy. This result is not entirely unexpected, as this test set had the largest domain shift. Another possible explanation for the poor performance on this test set is that the image selection algorithm may not be well suited to images from patients with heart conditions, as it may throw out useful images. A number of patients in this test set were left with significantly fewer images after the selection process compared to patients from the other test sets.

4. CONCLUSIONS

In this paper we demonstrated an online unsupervised learning method to boost performance of a COVID-19 classification model when tested with data with a domain shift from the training and validation sets. The aim of this method is to allow a model to adapt to a small domain shift in the data, without the need for expert labels.

Given the results of the experiments, we conclude that an online unsupervised learning method may be able to boost classification performance of COVID-19 diagnosis models under slight domain shift. However, further fine tuning is needed to see how much it can boost performance. Further explorations using different image selection algorithms may help boost performance on test set three. Testing on larger datasets may help clarify some of the current issues.

Table 2. Results of Experiments

	Results		
	Test set	***Model***	***Accuracy***
Exp. 1	Test Set 1	Baseline	0.9
Exp. 2	Test Set 2	Baseline	0.667
Exp. 3	Test Set 3	Baseline	0.633
Exp. 4	Test Set 1	Online Unsupervised	0.867
Exp. 5	Test Set 2	Online Unsupervised	0.767
Exp. 6	Test Set 3	Online Unsupervised	0.533

5. REFERENCES

[1] Parnian Afshar, Shahin Heidarian, Nastaran Enshaei, Farnoosh Naderkhani, Moezedin Javad Rafiee, Anastasia Oikonomou, Faranak Babaki Fard, Kaveh Samimi, Konstantinos N. Plataniotis, and Arash Mohammadi, "Covid-ct-md: Covid-19 computed tomography (ct) scan dataset applicable in machine learning and deep learning," 2020.

[2] Parnian Afshar, Shahin Heidarian, Farnoosh Naderkhani, Anastasia Oikonomou, Konstantinos N. Plataniotis, and Arash Mohammadi, "Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images," 2020.

[3] Shahin Heidarian, Parnian Afshar, Nastaran Enshaei, Farnoosh Naderkhani, Anastasia Oikonomou, S. Farokh Atashzar, Faranak Babaki Fard, Kaveh Samimi, Konstantinos N. Plataniotis, Arash Mohammadi, and Moezedin Javad Rafiee, "Covid-fact: A fully-automated capsule network-based framework for identification of covid-19 cases from chest ct scans," 2020.

[4] Xuehai He, Xingyi Yang, Shanghang Zhang, Jinyu Zhao, Yichen Zhang, Eric Xing, and Pengtao Xie, "Sample-efficient deep learning for covid-19 diagnosis based on ct scans," *medRxiv*, 2020.

[5] Richa Agarwal, Oliver Diaz, Xavier Lladó, Moi Hoon Yap, and Robert Martí, "Automatic mass detection in mammograms using deep convolutional neural networks," *Journal of Medical Imaging*, vol. 6, no. 3, pp. 1 – 9, 2019.

[6] Hiba Chougrad, Hamid Zouaki, and Omar Alheyane, "Convolutional neural networks for breast cancer screening: Transfer learning with exponential decay," *CoRR*, vol. abs/1711.10752, 2017.

[7] N. Khalifa M. Loey, G. Manogaran, "A deep transfer learning model with classical data augmentation and cgan to detect covid-19 from chest ct radiography digital images," 2020.

[8] Longlong Jing and Yingli Tian, "Self-supervised visual feature learning with deep neural networks: A survey," *CoRR*, vol. abs/1902.06162, 2019.

[9] Dina Ragab, Maha Sharkas, Stephen Marshall, and Jinchang Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," *PeerJ*, vol. 7, pp. e6201, 01 2019.

[10] L Chen, P Bentley, K Mori, K Misawa, M Fujiwara, and D Rueckert, "Self-supervised learning for medical image analysis using image context restoration," *Medical Image Analysis*, vol. 58, pp. 1–12, 2019.

[11] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick, "Momentum contrast for unsupervised visual representation learning," 2020.

[12] Francois Chollet, *Deep Learning with Python*, Manning Publications Co., USA, 1st edition, 2017.

[13] J. H. Moon, Debasmit Das, and C.S. George Lee, "Multi-step online unsupervised domain adaptation," *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020.

[14] Yuan Cao and Haibo He, "Learning from testing data: A new view of incremental semi-supervised learning," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 2008, pp. 2872–2878.

[15] Mohammad Rahimzadeh, Abolfazl Attar, and Seyed Mohammad Sakhaei, "A fully automated deep learning-based network for detecting covid-19 from a new and large lung ct scan dataset," *medRxiv*, 2020.

[16] Nicolas Ewen and Naimul Khan, "Targeted self supervision for classification on a small covid-19 ct scan dataset," 2020.

6. ACKNOWLEDGEMENTS

We acknowledge NSERC's funding through an Alliance grant to conduct this study.

BLIND DETECTION OF RADAR PULSE TRAINS VIA SELF-CONVOLUTION

Alex Byrley

Department of Electrical Engineering
University at Buffalo
Buffalo, NY 14260
Email: anbyrley@buffalo.edu

Adly Fam

Department of Electrical Engineering
University at Buffalo
Buffalo, NY 14260
Email: afam@buffalo.edu

ABSTRACT

This paper studies the blind detection of radar pulse trains using self-convolution. The self-convolution of a horizontally polarized pulse train with a constant pulse repetition frequency (PRF) is the same as its autocorrelation, only shifted in time, provided that the pulses are symmetric. This makes the waveform amenable to blind detection even in the presence of a constant Doppler shift. Once detected, we estimate the carrier, demodulate, and estimate the PRF of the baseband train using a logarithmic frequency domain matched filter. We derive a Neyman-Pearson self-convolution detection threshold for additive white Gaussian noise (AWGN) and conduct numerical experiments to compare the Signal-to-Noise Ratio (SNR) performance against standard matched filtering. We also illustrate the logarithmic frequency matched filter's PRF estimation accuracy.

Index Terms— radar, pulse-trains, blind-detection, self-convolution, prf-estimation, electronic-intelligence

1. INTRODUCTION

Blind detection is a fundamental task for autonomous electronic intelligence (ELINT) systems. A typical autonomous ELINT radar receiver passively detects and classifies an unknown number of signal emitters [1]. In doing so, we find that the constant PRF pulse train is ubiquitous within the electromagnetic spectrum due to its cost effective generation and use in legacy systems [2], despite the rapid advance in wireless waveform technology [3, 4]. Repeated motifs are also used in preambles [5, 6] as a mitigation against multipath and fading during autonomous device synchronization, thus accurate methods for blind detection of and parameter estimation for such pulse trains attracts considerable interest [7, 8, 9, 10, 11].

The self-convolution of horizontally polarized constant PRF pulse trains with symmetric pulses is the same as their autocorrelation, only shifted in time, enabling blind detection by an autononmous agent. This fact is due to the waveform's time domain symmetry [12]. A constant Doppler shift will not change this symmetry and thus self-convolution is Doppler shift invariant. Once detected, we demodulate the signal and estimate the PRF using a logarithmic frequency domain matched filter adapted from [13].

The blind detection problem is not new, as evidenced by the existence of complete textbooks [14] devoted to the subject, and neither is the exploitation of symmetry via self-convolution [15] nor PRF estimation [16]. This paper contributes a new detection and PRF estimation methodology designed for a special yet important case of a horizontally polarized constant PRF pulse train with symmetric pulses. We derive a novel self-convolution detection threshold for the scenario where we do not know the radar code nor the PRF of such a signal. We show that this applies even in the presence of a constant Doppler shift. The price of blind detection is a reduction in SNR performance relative to the case of having perfect knowledge of the signal and using a matched filter. Once detected, we estimate its PRF, which is a feature used for emitter classification [1], using a simple matched filter. The SNR ratio required for reliable detection is higher than that required for an accurate PRF estimation, and therefore, once detected, we are guaranteed an accurate PRF estimate.

Section 2 illustrates the self-convolution approach with an example. Section 3 describes the Neyman-Pearson detector. Section 4 details the logarithmic frequency matched filter. Section 5 shows the results of numerical experiments concerning the detection performance and PRF estimation accuracy. Section 6 concludes the paper.

2. BLIND DETECTION OF RADAR PULSE TRAINS

Due to the time domain symmetry of horizontally polarized constant PRF pulse trains with symmetric pulses, taking the self-convolution is equivalent to time shifting the autocorrelation. We show this via an illustrative example. Let $\mathfrak{z}(z) = z^0 + z^1 + z^2$ be the Z-transform of a constant PRF sequence. The autocorrelation is given by [17]:

$$R_{\mathfrak{z}\mathfrak{z}}(z) = \mathfrak{z}(z) \cdot \mathfrak{z}(z^{-1}) = z^{-2} + 2z^{-1} + 3 + 2z^1 + z^2 \quad (1)$$

In the same way, the self-convolution $S_{\mathfrak{z}\mathfrak{z}}(z)$ is given by:

$$S_{\mathfrak{z}\mathfrak{z}}(z) = \mathfrak{z}(z) \cdot \mathfrak{z}(z) = z^0 + 2z^1 + 3z^2 + 2z^3 + z^4 \quad (2)$$

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Therefore:

$$S_{\mathfrak{z}\mathfrak{z}}(z) = z^2 R_{\mathfrak{z}\mathfrak{z}}(z) \tag{3}$$

The Z-transform of an order N pulse train using a pulse $p(z)$ of non-zero width and experiencing a Doppler shift υ is expressed as $x(z) = p(z^\upsilon) \cdot \mathfrak{z}(z^\upsilon)$, where $\mathfrak{z}(\cdot)$ is the order N constant PRF sequence. Using the exact same technique as above, it is possible to show that in this case, we will have:

$$S_{xx}(z) = z^{\upsilon N} R_{pp}(z^\upsilon) R_{\mathfrak{z}\mathfrak{z}}(z^\upsilon) \tag{4}$$

provided that $p(z)$ is symmetric, i.e. $R_{pp}(z) = p(z) \cdot p(z^{-1}) = p(z) \cdot p(z) = S_{pp}(z)$. Thus the self-convolution of horizontally polarized constant PRF pulse trains with symmetric pulses is robust to constant Doppler shifts.

3. A BLIND DETECTOR FOR RADAR PULSE TRAINS

We suppose we receive a signal $x(t) = p(t)e^{j\omega_c t} * z(t) + w(t) = z_p(t) + w(t)$, where $z(t)$ is our pulse train signal with PRF f_{PRF} and symmetric pulses, $p(t)$ is the symmetric pulse, $*$ represents convolution, ω_c is the carrier frequency, and $w(t)$ is the AWGN. We assume $z_p(t)$ is horizontally polarized, so we take only the in-phase channel. We therefore have $w(t) \sim \mathcal{N}(0, \sigma^2)$, which is i.i.d.. The self-convolution of the signal after sampling is:

$$S_{xx}[n] = S_{z_p z_p}[n] + 2S_{z_p w}[n] + S_{ww}[n] \tag{5}$$

Our detector must decide between two hypothesis:

$$H_0 : \quad S_{xx}[n] = S_{ww}[n] \tag{6}$$

$$H_1 : \quad S_{xx}[n] = S_{z_p z_p}[n] + 2S_{z_p w}[n] + S_{ww}[n] \tag{7}$$

To derive the threshold, we first notice that if we assume our sampling rate is f_s, and we capture N pulses of the train $z(t)$, then our signal is of length $L = \frac{N f_s}{f_{prf}}$ after sampling. The noise self-convolution is therefore:

$$S_{ww}[n] = \sum_{m=0}^{L-1} w[m] \cdot w[n-m] \tag{8}$$

We can express this signal as a vector via the matrix multiplication:

$$\mathbf{S_{ww}} = \begin{bmatrix} w_0 & 0 & 0 & \cdots & 0 \\ w_1 & w_0 & 0 & \cdots & 0 \\ w_2 & w_1 & w_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & w_{L-1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{L-1} \end{bmatrix} = \mathbf{H}_w \mathbf{w} \tag{9}$$

then if we assume perfect sampling, i.e. $w[n] \sim \mathcal{N}(0, \sigma^2)$ are i.i.d., the elements of $S_{ww}[n]$ are inner-products of i.i.d. Gaussian random variables. If we make the additional assumption that L is even, which is not restrictive because the closest even number to $\frac{N f_s}{f_{prf}}$ is at most one sample away, the value of $S_{ww}[n]$ with the largest variance will be located at $n = \frac{L}{2}$, which corresponds to the row of $\mathbf{H}_w$ that has no zero elements. This means that our test statistic can be expressed as:

$$T_{|H_0} = \mathbf{w}_\mathrm{r}^T \mathbf{w} \tag{10}$$

where $\mathbf{w}_\mathrm{r}$ is this full row, which is just $w[-n]$, the reversed noise signal. Since L is even, denoting $M = \frac{L}{4}$ allows us to write the distribution of $T_{|H_0}$ as:

$$f_{T_{|H_0}}(x) = \frac{e^{-\frac{|x|}{2\sigma^2}}}{2\sigma^2(M-1)!} \cdot \sum_{k=0}^{M-1} \frac{(M+k-1)!}{2^{(M+k)} k!(M-k-1)!} \left(\frac{|x|}{2\sigma^2}\right)^{M-1-k} \tag{11}$$

The probability of false alarm is therefore given by:

$$P_{FA} = \frac{1}{(M-1)!} \sum_{k=0}^{M-1} \frac{(M+k-1)!}{2^{(M+k)} k!(M-k-1)!} \Gamma\left(M-k, \frac{\gamma}{2\sigma^2}\right) \tag{12}$$

where $\Gamma(M-k, \frac{\gamma}{2\sigma^2})$ is the upper incomplete gamma function. The detection threshold γ can be found via a simple line search since this expression is monotonic.

4. PRF ESTIMATION

After waveform detection via self-convolution, we estimate the carrier via the method in [18] and demodulate the signal to the baseband. We compute the logarithmic frequency power spectrum of the baseband train and then peform matched filtering. The maximum of the matched filter output will be at the PRF. The power spectrum of a baseband pulse train over the continuous time interval $[0, \frac{N}{f_{prf}}]$ is given by:

$$|X(f)|^2 = \frac{|P(f)|^2}{N} \cdot \frac{\sin^2\left(N\pi \frac{f}{f_{PRF}}\right)}{\sin^2\left(\pi \frac{f}{f_{PRF}}\right)} \tag{13}$$

where f_{PRF} is the PRF and $P(f)$ is the Fourier transform of the baseband pulse envelope. If we let $N \to \infty$, then we have:

$$|X(f)|^2 = |P(f)|^2 \cdot \sum_{k=-\infty}^{\infty} \delta\left(f - k f_{PRF}\right) \tag{14}$$

If we discard the non-positive frequencies, and we warp the remaining spectrum logarithmically, we get:

$$|X\left(\log(f)\right)|^2 = |P\left(\log(f)\right)|^2 \cdot \sum_{k=1}^{\infty} \delta\left(log(f) - log(k) - log(f_{PRF})\right) \tag{15}$$

We see that after sampling the logarithmic frequency domain, we can represent this warped spectrum as a sequence via the Z-transform:

$$|X(z)|^2 = |P(z)|^2 \cdot z^{-log(f_{PRF})} \sum_{k=1}^{\infty} z^{-log(k)} \tag{16}$$

If we therefore define a matched filter via:

$$M(z) = \sum_{k=1}^{\infty} z^{-log(k)} \tag{17}$$

then their crosscorrelation is:

$$R_{XM}(z) = |X(z)|^2 \cdot M^*(z^{-1}) \tag{18}$$

where $(\cdot)^*$ is complex conjugation. Using log-frequency notation, this is:

$$R_{XM}(log(f)) = |X(log(f))|^2 \cdot M^*(-log(f)) \tag{19}$$

Then since $P(f)$ is a lowpass function, we have

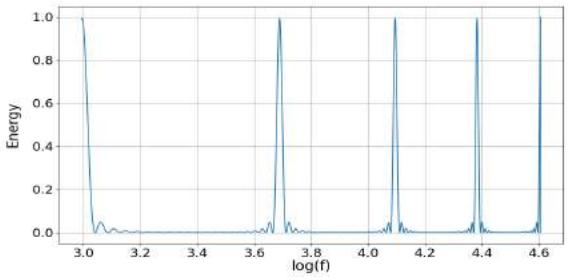
$$f_{PRF} = e^{\text{argmax}\{R_{XM}(log(f))\}} \tag{20}$$

Fig. 1 shows an example of using this matched filter on the logarithmic frequency power spectrum of a pulse train.

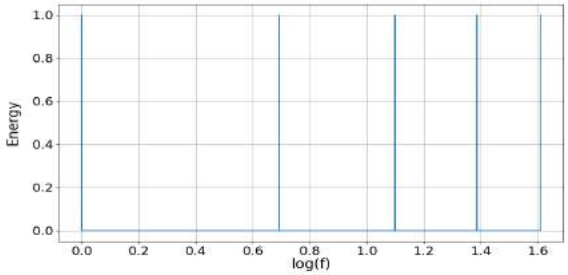
5. NUMERICAL EXPERIMENTS

5.1. Probability of Detection Comparison

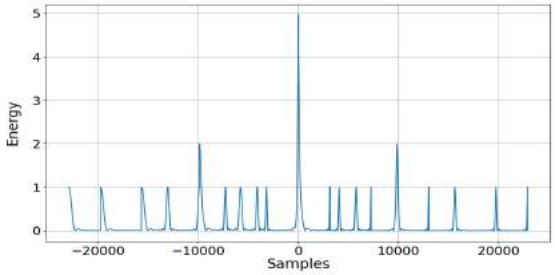
We compare the probability of detection (P_D) of a pulse train in AWGN using self-convolution with that of using a matched filter, which assumes have perfect knowledge of the signal. We use Monte Carlo to determine P_D due to the mutual dependence between $S_{z_p w}[n]$ and $S_{ww}[n]$ complicating the derivation of $f_{T_{|H_1}}(x)$. Both sets of detection results were obtained using 50,000 samples then fit with a sigmoid



(a) Pulse Train Logarithmic Frequency PSD $P(f) = 1 \; \forall f$

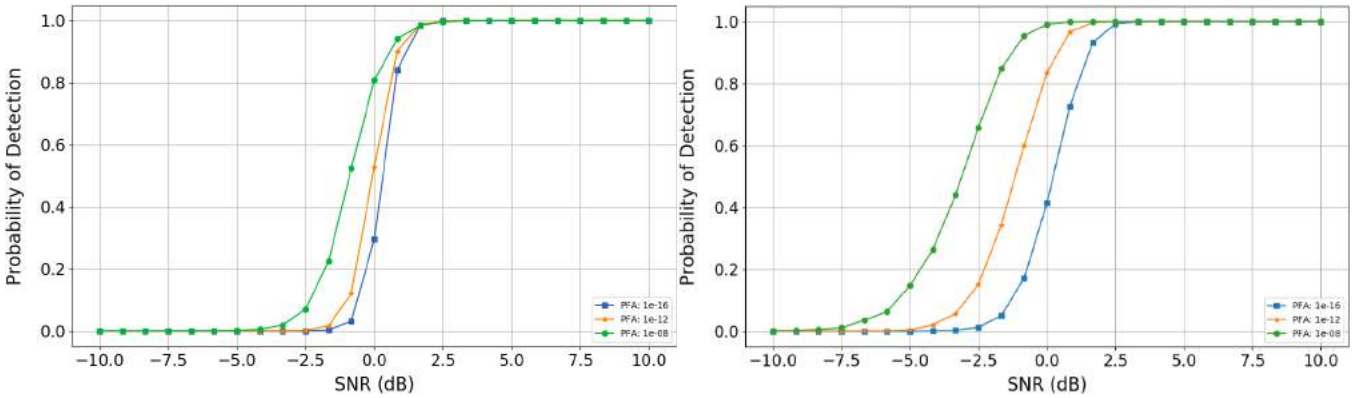


(b) Logarithmic Frequency Matched Filter



(c) Matched Filter Output $R_{XM}(log(f))$

Fig. 1: PRF Estimation Example



(a) Self-Convolution

(b) Matched Filter

Fig. 2: Probability of Detection Performance

curve. For simplicity of implementation, we used a simple rectangular pulse train with $L = 64$ and $N = \frac{L}{2}$.

Fig. 2 shows the results. We see that in exchange for blind detection and Doppler invariance, we lose SNR performance. This is to be expected since the matched filter is optimal in AWGN.

5.2. PRF Estimation Performance

Here we illustrate the performance of the PRF estimation procedure in AWGN. Assuming the pulse train has been detected and demodulated, we compute the percent PRF error for pulse trains of order $[5, 10, 15, 20]$. Each pulse train uses a square pulse and is contained in the interval $[0, 1]$ with $L = 2^{12}$ samples. Therefore an order N pulse train has N pulses in this interval. We use the Chirp Z-Transform to compute a critically sampled frequency transform over a bandwidth fives times greater than the PRF. Our modified matched filter has 100 elements, instead of the theoretically infinite number. We perform 1000 iterations for each SNR.

To put our method in context, we compare its performance against finding the first peak of the incoming signal's autocorrelation after filtering it with a maxflat filter [19] for noise reduction. The filter had a normalized cutoff frequency of $\omega_c = 0.034$ and used 256 zeros. Maxflat filters minimize spectral amplitude distortions in their passbands and so we choose them to preserve the autocorrelation peaks while achieving noise reduction. Fig. 2 shows the mean and standard deviation of the absolute value of the percent error of the PRF estimate averaged across order for each SNR where detection is likely according to Fig. 2. Fig. 2 shows that our proposed method should be preferred for SNRs < 0dB, that the methods are roughly equivalent considering the standard deviation at an SNR of 0dB, and the simple noise filtering method is better above 0dB.

We note that choosing the cutoff frequency ω_c can be difficult in practice because the number of pulses N and the length of the signal L is unknown a-priori and therefore any chosen ω_c runs the risk of destroying the periodicity in the spectrum. Fig.2 shows that the proposed PRF estimation procedure provides substantial noise reduction benefits, as if a noise reduction filter had been appropriately designed, without having to know the needed signal information to design such a filter. In other words, the proposed approach offers similar noise reduction benefits as if we had perfect information about the signal and could design a strong noise reduction filter that would not harm the signal's periodicity.

6. CONCLUSION

This paper presented a method for blind detection and PRF estimation of constant PRF radar pulse trains with symmetric pulses. The self-convolution of such a pulse train is the same as its autocorrelation, only shifted in time, and therefore the

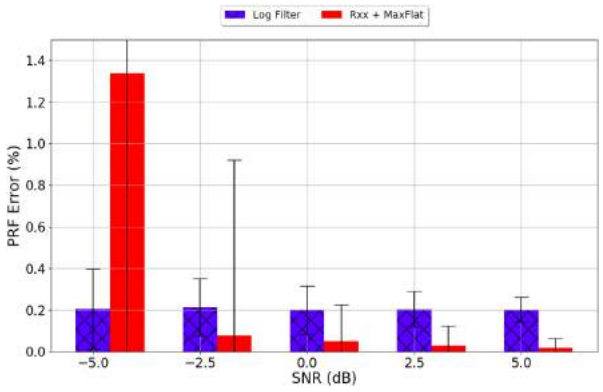


Fig. 2: Percent Error Comparsion

waveform is amenable to blind detection. This observation provides an important mechanism for autonomous systems, such as UAVs running ELINT missions, to passively detect hostile emitters.

We derived a Neyman-Pearson self-convolution detection threshold for additive white Gaussian noise and showed that it does not perform as well as the matched filter. However, in exchange for a higher SNR requirement, we get blind detection and Doppler invariance. Thus, we do not have to know the code, nor its PRF, yet we can still detect the waveform at the price of reduced noise tolerance with respect to what would be required if we knew both the code and the PRF.

Once our autonomous system has blindly detected an emitter's waveform, we use a logarithmic frequency matched filter to estimate the emitter's PRF. We showed that our PRF estimation technique is capable of estimating the PRF to within one percent for SNRs in which self-convolution has a chance of passive detection. We showed that this general level of accuracy is also achievable by passing the received signal's autocorrelation through a maxflat filter to remove the noise. The design of such a filter however requires information about the signal's periodicity, to which we are not privy by virtue of the blind detection problem. Our method for blind PRF estimation therefore provides powerful noise reduction capabilities without needing the information required to design a noise reduction filter and thereby showing it is a true blind PRF estimation technique.

Future work will be focused on extending these results to more general pulse trains, such as pulse trains with missing samples, jittered pulse trains, and trains with non-symmetric pulse shapes. Symmetric pulse trains with missing samples are still detectable via self-convolution, however they will provide a lower probability of detection by virtue of having less signal power. Finally, we note that the magnitude of the analytic signal of a constant amplitude non-symmetric pulse shape, such as a chirp, will be symmmetric, and therefore trains using these pulses can still be blindly detected via self-convolution provided we compute the analytic signal via [20] first.

7. REFERENCES

[1] R. G. Wiley, *Electronic Intelligence: The Analysis of Radar Signals*. Norwell, MA: Artech House, 2006.

[2] R. A. Romero and K. D. Shepherd, "Friendly spectrally shaped radar waveform with legacy communication systems for shared access and spectrum management," *IEEE Access*, vol. 3, pp. 1541–1554, 2015.

[3] Z. Liu, "Recognition of multi-function radars via hierarchically mining and exploiting pulse group patterns," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–1, 2020.

[4] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[5] Q. Zhou, S. Wu, and D. Yu, "Preamble detection based on repeated preamble codes," U.S. Patent 8 599 569, Oct. 15, 2013.

[6] D. I. Im and Y. S. Seo, "Duty cycle detection circuit and semiconductor apparatus including the same," U.S. Patent 9 537 490, Jan. 3, 2017.

[7] T. L. Conroy and J. B. Moore, "On the estimation of interleaved pulse train phases," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3420–3425, 2000.

[8] Chen Yingying, Yang Jiankang, and Huo Jinghe, "Estimation algorithm of prf for space-borne sar based on local wave decomposition(lwd) pulse compression," in *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2016, pp. 786–789.

[9] U. I. Ahmed, T. ur Rehman, S. Baqar, I. Hussain, and M. Adnan, "Robust pulse repetition interval (pri) classification scheme under complex multi emitter scenario," in *2018 22nd International Microwave and Radar Conference (MIKON)*, 2018, pp. 597–600.

[10] D. Benvenuti, "Genetic algorithms for pri ambiguity resolution in passive emitter tracking," in *2009 European Radar Conference (EuRAD)*, 2009, pp. 117–120.

[11] S. Wu, W. Su, L. Zhu, Y. Wang, and X. Shan, "Algorithm based on pri transform for estimating the scanning periods of phased array radar," in *2009 5th International Conference on Wireless Communications, Networking and Mobile Computing*, 2009, pp. 1–4.

[12] F. A. Qazi and A. T. Fam, "Interception of the triangular fm waveform via self-convolution," in *MILCOM 2015 - 2015 IEEE Military Communications Conference*, 2015, pp. 1515–1518.

[13] A. N. Byrley, "Logarithmic frequency waveforms: A new paradigm in radar signal processing," Ph.D. dissertation, The State University of New York at Buffalo, 2020.

[14] X. Shi, *Blind Signal Processing: Theory and Practice*. Berlin: Springer, 2012.

[15] Y. T. Chan, B. H. Lee, R. Inkol, and F. Chan, "Estimation of pulse parameters by autoconvolution and least squares," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 363–374, 2010.

[16] R. M. Hawkes, "Radar emitter recognition using pulse repetition interval," in *In Proceedings of the International Conference on Information Systems and Science*, 1979.

[17] A. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Upper Saddle River, New Jersey: Prentice-Hall Inc., 1999.

[18] A. T. Fam and R. Kadlimatti, "Blind interception of phase coded signals," in *2016 IEEE Radar Conference (RadarConf)*, 2016, pp. 1–5.

[19] L. Rajagpoal and S. D. Roy, "Design of maximally-flat fir filters using the bernstein polynomial," *IEEE Transactions on Circuits and Systems*, vol. 34, no. 12, pp. 1587–1590, 1987.

[20] D. Romero and G. Dolecek, "Digital fir hilbert transformers: Fundamentals and efficient design methods," in *MATLAB: A Fundamental Tool for Scientific Computing and Engineering Applications - Volume 1*, V. Katsikis, Ed. London, UK: Intech Open Ltd., 2012, pp. 445–482.

Order Dispatching in Ride-Sharing Platform under Travel Time Uncertainty: A Data-Driven Robust Optimization Approach

Xiaoming Li [1], *Student Member, IEEE,* Jie Gao [2], Chun Wang [3], *Member, IEEE,* Xiao Huang [4], Yimin Nie [5]

***Abstract*— In this paper, we study a one-to-one matching ride-sharing problem to save the travellers' total travel time considering travel time uncertainty. Unlike the existing work where the uncertainty set is assumed to be known or roughly estimated, in this work, we propose a learning-based robust optimization framework to handle the issue properly. Specifically, we assume the travel time varies in an uncertainty set which is predicted by a machine learning approach - ARIMA using travel time historical data, the predicted uncertainty set then serves as the input parameter for the robust optimization model. To evaluate the proposed approach, we conduct a group of numerical experiments based on New York taxi trip record data sets. The results show that our proposed data-driven robust optimization approach outperforms the robust optimization model with a given uncertainty set in terms of total travel time savings. Further, the proposed approach can improve the travel time savings up to 112.8%, and 34% by average. Most importantly, our proposed approach is capable of handling the uncertainty in a more effective way when the uncertainty degrees become high.**

***Index Terms*— Data-Driven Optimization, Robust Optimization, Ride-Sharing, Travel Time Savings**

I. Introduction

Leveraging on the widely used network and the advances in the intelligent transportation systems (ITS), the large-scale ride-sharing platforms such as Didi, Uber and Lyft have dramatically reshaped the passengers' travel pattern. Ride-sharing systems aim to match riders with similar itineraries to reduce the number of vehicles to alleviate traffic congestion and save energy consumption. On the one hand, people may benefit from ride-sharing service by sharing to reduce the travel cost which makes the Mobility on Demand (MoD) more popular, on the other hand, a huge volume of traffic data generates in MoD systems which bring more challenges for ride-sharing system. One of the key issues for the ride-sharing platform is how could it quickly and effectively match drivers and riders in real-time under data-driven environment. In this paper, we study a dynamic ride-sharing travel time savings problem where a large number of drivers and riders are matched under travel time uncertainty. To be specific, we consider one-to-one driver and rider matching for the sake of simplicity. Drivers depart from their origins to pick up the assigned riders once the riders send their requests. The matched driver will travel to the assigned rider's origin to served the rider and drop the rider at his/her destination to finish the ride-sharing service. The objective of the problem is to maximize the travel time savings under travel time uncertainty.

Ride-sharing problems have attracted many researchers in the field of operational research (OR) and intelligent transportation systems (ITS). In this paper, we focus on the work that is related to ride-sharing modeling technique. Armant et al. [1] study ride-sharing systems to reduce the number of vehicles driving on roads. To minimize the total rider trip distance, a mixed integer programming is proposed in this work. Duan et al. [2] consider both drivers' idle driving distances and riders' waiting time reduction simultaneously. They propose a dynamic programming algorithm to solve the increase ratio of the size of the dispatch regions where rider requests are sent to the drivers in the dispatch region chosen by the central system. Li et al. [3] investigate the ride-sharing systems considering travel time uncertainty, the objective is to minimize the overall cost of the entire system. Since the travel time of shared vehicles has a great impact on the optimization solution, the problem is formulated using robust optimization [4] modeling technique. Guo et al. [5] propose a real-time ride-sharing framework with dynamic time window to optimize large-scale online ride-sharing platform. A multi-strategy solution graph search heuristic approach is proposed to derive high-quality solutions. Lin et al. [6] explore the ride-sharing routing problem considering travel demand uncertainty. The problem is formulated as a two-stage stochastic optimization problem and solved by a demand-aware approach. Wang et al. [7] study several stable matching strategies for dynamic ride-sharing systems where greedy matching method, basic stable formulation and nearly stable formulation are discussed. In addition, a deterministic rolling horizon framework is proposed to

[1]Xiaoming Li is with the Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montréal, QC H3G 1M8, Canada, with the school of computer, Shenyang Aerospace University, Shenyang, 110136, China, and also with the Global Artificial Intelligence Accelerator (GAIA) innovation hub, Ericsson INC., Montréal, QC H4R 2A4, Canada xiaoming.li@mail.concordia.ca

[2]Jie Gao is with the Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montréal, QC H3G 1M8, Canada, and also with the Global Artificial Intelligence Accelerator (GAIA) innovation hub, Ericsson INC., Montréal, QC H4R 2A4, Canada jie.gao@mail.concordia.ca

[3]Chun Wang is with the Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montréal, QC H3G 1M8, Canada chun.wang@concordia.ca

[4]Xiao Huang is with the Concordia John Molson School of Business (JMSB), Concordia University, Montréal, QC H3G 1M8, Canada xiao.huang@concordia.ca

[5]Yimin Nie is with the Global Artificial Intelligence Accelerator (GAIA) innovation hub, Ericsson INC., Montréal, QC H4R 2A4, Canada yimin.nie@ericsson.com

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

handle the uncertainty in ride-sharing systems. Wang et al. [8] design a real-time high-capacity ride-sharing model with subsequent information. Since riders may not be served at the current time for those high-demand and high-density regions, this paper applies subsequent information to pursue more optimal route arrangements for riders. Ma et al. [9] propose a ride-sharing strategy with integrated transit. Additionally, vehicle dispatch and idle vehicle relocation algorithm based on queue theory are customized for the problem. The objective is to minimize rider access time tot idle vehicles and total idle vehicle relocation cost. Sun et al. [10] propose an optimal demand-responsive transit (DRT) model to transport passengers of all demand points to the transportation hub such as railway, airport etc. The objective is tot minimize the weighted passenger walking and riding time simultaneously. Farhan el al. [11] focus on the ride-sharing problem with shared autonomous electric vehicle (SAEV). After examining the charging technology and charging infrastructure, this work integrates one-way ride-sharing to determine the impact of SAEV which includes identifying fleet size, charging infrastructure sites and riders' waiting time. Jain et al. [12] formulate a weighted graphing coloring optimization problem to promote ride-sharing efficiency to maximize the overall revenue. Additionally, a ride-sharing framework which considers quality of sharing and operator revenue is proposed. Simonetto el al. [13] propose a novel solution for the city-scale ride-sharing problem to maximize computational efficiency. The solution combines a linear assignment algorithm, a context-mapping algorithm and a capacitated vehicle routing problem (CVRP) with pick-up, delivery and time-window.

For ride-sharing problems, most existing work adopt deterministic model to formulate the problem, in other words, all the parameters in the optimization models are fixed without any uncertainty. However, in the real ride-sharing applications, optimization parameters such as rider demand, driver travel time are subject to uncertainty. Although [3] and [6] consider modeling uncertainty via robust optimization and stochastic optimization [14], the parameters of the uncertainty set and probability distribution are assumed to be known. More specifically, [3] assumes the driver travel time vary in a given range, and [6] assumes the probability distribution type is given. Such assumptions may lead the stochastic or robust models sub-optimal or even infeasible [15]. To fill the research gap to make the robust optimization models more practical for ride-sharing travel time savings problem under data-driven environment, we propose a learning-based robust optimization framework which integrates machine learning method and robust optimization modeling technique to tackle the travel time uncertainty.

The rest of this paper is organized as follows. The ride-sharing travel time savings problem studied in this paper is described in Section II. The data-driven robust optimization framework is presented in Section III. In Section IV, we validate our proposed approach through a group of numerical experiments based on New York taxi trip record, finally we conclude our work and provide extensions of the proposed approach in Section V.

II. Problem Statement

We consider a ride-sharing platform which consists of a central service operator, a set of drivers and a set of riders. Drivers and riders enter the platform and exit when they share a ride, reach their destinations, or meet their deadlines. The goal of the service operator is to dynamically match drivers to riders such that the total travel time savings is maximized by considering the travel time uncertainty. We assume that only one rider can be assigned to (i.e., share a ride with) one driver, however, the proposed model can generally be extended for assigning multiple riders to a single driver by adding several simple constraints.

Let $\mathcal{A}^k$ denote a set of travellers (drivers or riders) that are in the platform at time k. Let $\mathcal{L}$ be the set of locations. Each traveller $a \in \mathcal{A}^k$ is characterized by a trip request $\langle o(a), w(a), es(a), la(a)\rangle$, where $o(a) \in \mathcal{L}$ and $w(a) \in \mathcal{L}$ are the origin and destination locations of a, $es(a)$ and $la(a)$ represent the earliest departure time and latest arrival time of a at time k, respectively. The set $\mathcal{A}^k$ of travellers in the platform at time t is partitioned into two sets: a set of drivers $\mathcal{D}^k$ and a set of riders $\mathcal{R}^k$. Then each driver and rider can be characterized by the trip request $\langle o(d), w(d), es(d), la(d)\rangle$ and $\langle o(r), w(r), es(r), la(r)\rangle$, respectively. Let dt_d be the departure time of driver d. Let $x_{d,r} = 1$ if driver $d \in \mathcal{D}^k$ is matched with rider $r \in \mathcal{R}^k$, and 0 otherwise. One necessary condition of the ride assignment is that the travel distance savings which is defined in (1) must be positive. In this work, we assume the travel time is proportional to the travel distance, hence, the necessary condition becomes the travel time savings is positive which reflects in the first constraint of ride-sharing robust optimization model.

$$dist(o(d), w(d)) - dist(o(d), o(r)) - dist(w(r), w(d)) \quad (1)$$

Travel time uncertainty can play a critical role in our ride-sharing problem, as the optimal solution could be even infeasible (some time window constraints may be violated) when realized travel time is significantly different than its nominal value. Let $t_{i,j}$ be the realized travel time between two locations $i \in \mathcal{L}$ and $j \in \mathcal{L}$, with $t_{i,j} = \bar{t}_{i,j} + \xi_{i,j}\tilde{t}_{i,j}$. For here $\bar{t}_{i,j}$ is the nominal travel time and $\tilde{t}_{i,j}$ is the travel time deviation between locations i and j. The basic process of the one-to-one ride-sharing service discussed in this paper is illustrated in Fig. 1.

III. The Data-Driven Robust Optimization Framework

The proposed data-driven robust optimization framework involves three steps, namely, learning strategy for uncertainty set construction, one-stage robust optimization modeling, and robust counterpart reformulation. We will discuss the holistic learning-based framework in details in this section.

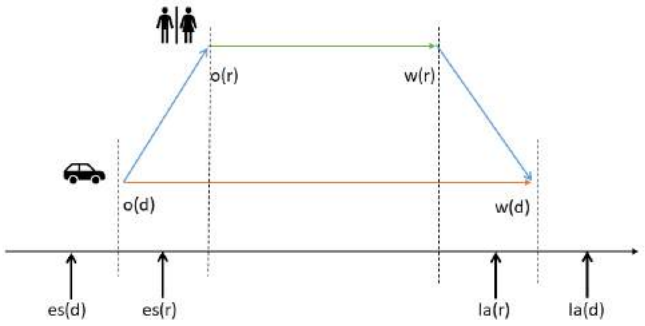


Fig. 1: The process of one-to-one matching ride-sharing service. The orange line and green line are the driver and rider's trip routes, respectively without ride-sharing service. The two blue lines denote the rider's pick up route and the driver's route after dropping off the matched rider, respectively. The entire ride-sharing route is $o(d) \longrightarrow o(r) \longrightarrow w(r) \longrightarrow w(d)$.

TABLE I: Notation Table for Mathematical Models

Sets	**Description**
$\mathcal{D}^k$	A set of drivers at time k, indexed by d
$\mathcal{R}^k$	A set of riders at time k, indexed by r
$\mathcal{L}$	A set of regions indexed by i and j
$\mathcal{A}^k$	A set of travelers at time k, indexed by a, $\mathcal{A}^k = \mathcal{D}^k \cup \mathcal{R}^k$
$\mathcal{T}$	A set of time slots, indexed by k
$\mathcal{U}$	The uncertainty set of travel time between regions
Parameters	**Description**
$t_{i,j}$	The realized travel time from region i to j, $t_{i,j} = \bar{t}_{i,j} + \xi_{i,j}\tilde{t}_{i,j}$
$\bar{t}_{i,j}$	The nominal travel time from region i to j
$\tilde{t}_{i,j}$	The travel time deviation from region i to j
Γ	The uncertainty degree of the polyhedral uncertainty set
$o(a)$, $w(a)$	The origin and destination of a, $o(a), w(a) \in \mathcal{L}$
$es(a), la(a)$	The earliest starting time and latest arrival time of a)
Variables	**Description**
$x_{d,r} \in \{0,1\}$	Matching status that is equal to 1 if driver d and rider r is matched
$dt_d \in \mathbb{R}_+$	The departure time of driver d
$\xi_{i,j} \in \mathcal{U}$	Random variables whose values vary in the given uncertainty sets

A. Uncertainty Set Construction

A few work regarding to the strategy of uncertainty set construction is discussed in literature. For example, Dirichlet process mixtures and variational inference algorithm is introduced to construct uncertainty set based on labeled data [16]. The approach cannot be applied in our problem since the travel time in ride-sharing is a typical time-series data. For a better travel time prediction, in this work, we use AutoRegressive Integrated Moving Average (ARIMA) [17] algorithm to predict the nominal and maximum deviation of travel times between the regions based on the travel time historical data. Given a time-series sequence of historical travel time data $(t^1_{i,j}, t^2_{i,j}, \cdots, t^n_{i,j})$, the predicted nominal travel time and deviation of travel time can be derived by AIRMA which denote as $\bar{t}^{n+1}_{i,j}$ and $\tilde{t}^{n+1}_{i,j}$, respectively. Thereafter, the realized travel time takes values in $\left[\bar{t}^{n+1}_{i,j}, \bar{t}^{n+1}_{i,j} + \tilde{t}^{n+1}_{i,j}\right]$ which is treated as the uncertainty set.

B. Robust Optimization Model

With the objective of maximizing the total travel time savings of all matched drivers and riders, the ride-sharing optimization problem is formulated as following robust optimization model.

$$\max \quad \sum_{d\in\mathcal{D}^k}\sum_{r\in\mathcal{R}^k}\left(\overline{T}^0_{d,r}x_{d,r} + \min_{\xi\in\mathcal{U}} \xi^0_{d,r}\widetilde{T}^0_{d,r}x_{d,r}\right) \tag{2}$$

s.t.

$$\overline{T}^0_{d,r} + \min_{\xi\in\mathcal{U}} \xi^0_{d,r}\widetilde{T}^0_{d,r} + H(1-x_{d,r}) \geqslant 0, \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{2a}$$

$$dt_d \geqslant es(d), \quad \forall d \in \mathcal{D}^k, \tag{2b}$$

$$dt_d + \bar{t}_{o(d),o(r)} + \min_{\xi\in\mathcal{U}} \xi_{o(d),o(r)}\tilde{t}_{o(d),o(r)} + H(1-x_{d,r}) \geqslant es(r), \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{2c}$$

$$dt_d + \overline{T}^1_{d,r} + \min_{\xi\in\mathcal{U}} \xi^1_{d,r}\widetilde{T}^1_{d,r} \leqslant la(r) + H(1-x_{d,r}), \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{2d}$$

$$dt_d + \overline{T}^2_{d,r} + \min_{\xi\in\mathcal{U}} \xi^2_{d,r}\widetilde{T}^2_{d,r} \leqslant la(d) + H(1-x_{d,r}), \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{2e}$$

$$\sum_{r\in\mathcal{R}} x_{d,r} \leqslant 1, \quad \forall d \in \mathcal{D}^k, \tag{2f}$$

$$\sum_{d\in\mathcal{D}} x_{d,r} \leqslant 1, \quad \forall r \in \mathcal{R}^k, \tag{2g}$$

$$x_{d,r} \in \{0,1\}, \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{2h}$$

$$dt_d \in \mathbb{R}_+, \quad \forall d \in \mathcal{D}^k, \tag{2i}$$

$$\overline{T}^0_{d,r} = \left(\bar{t}_{o(d),w(d)} - \bar{t}_{o(d),o(r)} - \bar{t}_{w(r),w(d)}\right), \tag{2j}$$

$$\widetilde{T}^0_{d,r} = \left(\tilde{t}_{o(d),w(d)} - \tilde{t}_{o(d),o(r)} - \tilde{t}_{w(r),w(d)}\right), \tag{2k}$$

$$\xi^0_{d,r} = \left(\xi_{o(d),w(d)} - \xi_{o(d),o(r)} - \xi_{w(r),w(d)}\right), \tag{2l}$$

$$\overline{T}^1_{d,r} = \left(\bar{t}_{o(d),o(r)} + \bar{t}_{o(r),w(r)}\right), \tag{2m}$$

$$\widetilde{T}^1_{d,r} = \left(\tilde{t}_{o(d),o(r)} + \tilde{t}_{o(r),w(r)}\right), \tag{2n}$$

$$\xi^1_{d,r} = \left(\xi_{o(d),o(r)} + \xi_{o(r),w(r)}\right), \tag{2o}$$

$$\overline{T}^2_{d,r} = \left(\bar{t}_{o(d),o(r)} + \bar{t}_{o(r),w(r)} + \bar{t}_{w(r),w(d)}\right), \tag{2p}$$

$$\widetilde{T}^2_{d,r} = \left(\tilde{t}_{o(d),o(r)} + \tilde{t}_{o(r),w(r)} + \tilde{t}_{w(r),w(d)}\right), \tag{2q}$$

$$\xi^2_{d,r} = \left(\xi_{o(d),o(r)} + \xi_{o(r),w(r)} + \xi_{w(r),w(d)}\right). \tag{2r}$$

The objective (2) is to maximize the total travel time savings of matched riders and drivers under the travel time worst case scenario. We define a group of short-hand expressions (2j) - (2r) for the sake of succinctness. Constraints (2a) ensure that the travel time savings must be positive if driver d is assigned to rider r, (2b) ensure that the departure time of driver d must be later than his/her earliest starting time, constraints (2c) guarantee that the rider's drop off time must be earlier than his/her latest arrival time if driver d is assigned to rider r, constraints (2d) imply that the drop off time for rider j must be earlier than the rider j's latest arrival time, constraints (2e) imply that driver d must be able to reach his/her destination before his/her latest arrival time if rider j if rider j's if picked up, constraints (2f) and (2g) enforce that driver and rider are one-to-one matched in this model, constraints (2h), (2i) specify the types decision variables.

C. The Robust Counterparts

Since RO models cannot be solved by any off-the-shelf mathematical solvers, one possible way is to utilize robust couterpart (RC) which can convert RO models to deterministic models. We consider using the polyhedral uncertainty sets [4] which most robust optimization work considers.

1) Polyhedral Uncertainty Set (1-norm):

$$U_1 = \{\xi \mid \|\xi\|_1 \leq \Gamma\} = \left\{\xi \mid \sum_{j\in J_i} |\xi_j| \leq \Gamma\right\} \tag{3}$$

Given a set of $x^*_{d,r}$ values, the second term of the objective function (2), namely, $\min_{\xi\in\mathcal{U}} \xi^0_{d,r}\widetilde{T}^0_{d,r}x_{d,r}$ is equivalent to the following optimization problem.

$$\min \quad \widetilde{T}^0_{d,r}x^*_{d,r}z_{d,r} \tag{4}$$

s.t.

$$\sum_{d\in\mathcal{D}^k}\sum_{r\in\mathcal{R}^k} z_{d,r} \geqslant \Gamma, \tag{4a}$$

$$z_{d,r} \geqslant 1, \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k. \tag{4b}$$

Then we come to the following theorem.

Theorem 1. *The RC of the ride-sharing robust model has the equivalent formulation as follows.*

$$\max \quad \sum_{d\in\mathcal{D}^k}\sum_{r\in\mathcal{R}^k}\left(\overline{T}^0_{d,r}x_{d,r} + \Gamma g + h_{d,r}\right) \tag{5}$$

s.t. (2b), (2f) - (2i)

$$g + h_{d,r} \leqslant \Gamma\widetilde{T}_{d,r}x_{d,r}, \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{5a}$$

$$\overline{T}^0_{d,r} + \Gamma\widetilde{T}^o_{d,r} + H(1 - x_{d,r}) \geqslant 0, \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{5b}$$

$$dt_d + \bar{t}_{o(d),o(r)} + \Gamma\tilde{t}_{o(d),o(r)} \geqslant es_r + H(1 - x_{d,r}), \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{5c}$$

$$dt_d + \overline{T}^1_{d,r} + \Gamma\widetilde{T}^1_{d,r} \leqslant la_r + H(1 - x_{d,r}), \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{5d}$$

$$dt_d + \overline{T}^2_{d,r} + \Gamma\widetilde{T}^2_{d,r} \leqslant la_d + H(1 - x_{d,r}), \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k. \tag{5e}$$

$$g \in \mathbb{R}_+, \tag{5f}$$

$$h_{d,r} \in \mathbb{R}_+, \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k. \tag{5g}$$

where g and h are dual variables.

Proof: The dual problem of (4) can be written as follows:

$$\max \quad \Gamma g + \sum_{d\in\mathcal{D}^k}\sum_{r\in\mathcal{R}^k} h_{d,r} \tag{6a}$$

$$g + h_{d,r} \leqslant \Gamma\widetilde{T}^0_{d,r}x_{d,r}, \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k, \tag{6b}$$

$$g \in \mathbb{R}_+, \tag{6c}$$

$$h_{d,r} \in \mathbb{R}_+, \quad \forall d \in \mathcal{D}^k, \forall r \in \mathcal{R}^k. \tag{6d}$$

IV. Numerical Experiments

In this section, we evaluate our proposed data-driven robust optimization model through a group of simulations based on the real data sets. We use New York taxi trip records[1] from January 2017 to June 2017 for the machine learning model training and testing. The uncertainty set construction using ARIMA is implemented by Python 3.7, the reformulated RO models are solved by the commercial solver Gurobi 9.0[2]. The experiments are run on a PC with Intel Core i7 CPU, 32GB RAM, Windows 10.

A. Data Process

The New York taxi trip data sets consist of pick up/drop off locations and pick up/drop off times of each rider request. The data sets are split into training set (January 2017 to May

[1]https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

[2]https://www.gurobi.com/academia/academic-program-and-licenses/

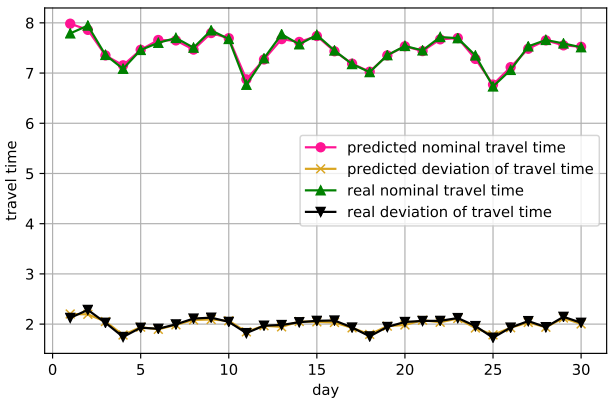


Fig. 2: The comparisons of average predicted nominal & deviation of travel time and real nominal & deviation of travel time (in minutes) in June 2017

2017) and testing set (June 2017). In the trip records, the pick up and drop off points fall into 256 region, for the ride-sharing purpose, we select 7 locations with the location ids 41, 42, 74, 75, 166, 236 and 238 where the rider demands are relatively higher. The locations information details can be found via https://data.world/nyc-taxi-limo/taxi-zone-lookup.

Additionally, we remove the rider trip records whose trip duration (difference of rider's drop off time and pick up time) are less than 1 minute and greater than 15 minutes for the sake of real ride-sharing application scenario. We discretize one day into 24 time windows (1 hr for each time window), then we compute the mean and standard deviation of the travel times between regions as the nominal and deviation of travel times during each time slot, respectively. Finally, we obtain the sequence of nominal and deviation of travel times for each paired of regions.

B. Experiment Results

We adopt ARIMA to predict the nominal and deviation of travel times for the paired of regions in June. The comparative results of the nominal travel time and deviation of travel time are shown in Fig. 2.

For data-driven robust optimization results, we choose 6 time slots (June 1st 0am and 17pm in weekday, June 3rd 0am and 17pm in weekend, June 18th 0am and 17 am for Father's Day). The number of riders can be easily calculated by trip record aggregation. Meanwhile, since there is no driver information in the given data sets due to the privacy issue, we random generate drivers across regions for the simulation. We assume the number of drivers in each time slot is 20% more than the number of riders without loss of generality. The numbers of riders and drivers in these time slots are listed in TABLE II.

The comparisons of total and average travel time savings by our proposed data-driven robust optimization approach and regular robust optimization are shown in TALBE III and IV, respectively. It is observed that the travel time savings by ride-sharing decrease as the uncertainty degree

TABLE II: The total number of riders and drivers in the ride-sharing regions during the given time slots

Time Slots No.	1	2	3	4	5	6
Num. of Rider	33	209	102	158	67	153
Num. of Drivers	40	251	123	190	81	184

becomes high since the worst-case solution is generated, the trend is consistent through all the 6 different time slots. Notice that if $\Gamma = 0$, the RO model becomes the corresponding deterministic (nominal) model without any uncertainty which yields the highest value but the solution fails to hedge against the uncertainty (constraints violation). Furthermore, our proposed data-driven robust optimization approach consistently outperforms the roust model without learning method across the 6 time slots and the uncertainty degrees (values of Γ) in terms of travel time savings, the largest gap can reach up to 112.8%, and the average gap is 34%. In particular, the gaps between data-driven and regular robust optimization dramatically increase when Γ becomes high which indicates that our proposed approach is capable of capturing the uncertainty more effectively.

Additionally, we compare the data-driven robust optimization approach with the robust optimization given a predefined uncertainty set in terms of the violation rates. In this work, the *violation rate* is defined as the number of matched rider divided by the number of requested rider. The comparative results are shown in Fig. 3. It is observed that the data-driven approach (labeled by dd-time slot) is lower than the non-driven approach (labeled by non-dd-time slot) in terms of violation rate (X-axis), the trend is fairly consistent under the levels of uncertainty degree across the three days.

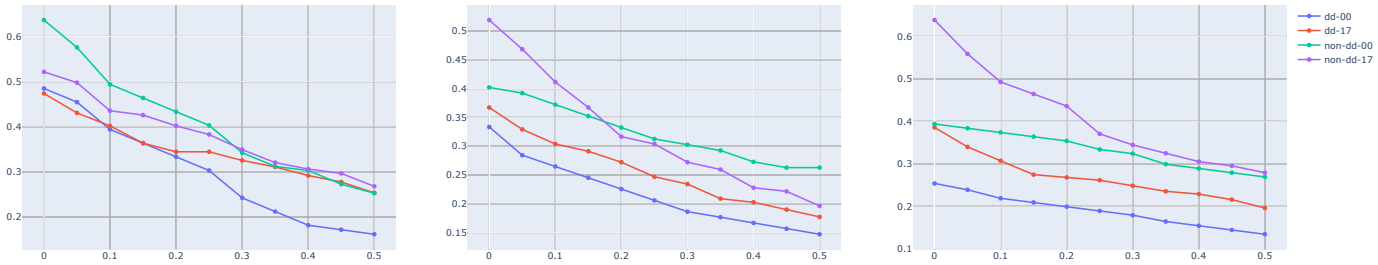
V. Conclusions and Future Work

In this paper, we propose a novel data-driven robust optimization approach to address the ride-sharing travel time savings problem under travel time uncertainty. The approach integrates a learning approach ARIMA with robust optimization modeling technique which can be applied to other homogeneous applications not limited to ride-sharing problem. In reality, our proposed data-driven optimization framework could be easily extended by the components replacement. For example, different machine/deep learning algorithms can be adopted to derive the appropriate parameters for optimization models from historical data. Also, other optimization under uncertainty techniques such as stochastic programming [18], [19] and distributionally robust optimization [20] can be used for problem modeling. In addition, there are several potential research directions that are worth to follow up.

- The ride-sharing travel time savings problem in this paper is formulated as a one-stage robust optimization model whose solutions may be conservative in some applications, we plan to utilize adaptive (two-stage) robust optimization modeling to make the ride-sharing problem more practical.

TABLE III: The comparison of total travel time savings for the ride-sharing systems (in minutes) by data-driven and non-data-driven robust optimization under different levels of uncertainty degree. For each time slot, the top row is derived from data-driven robust optimization, and the bottom row is derived from non-data-driven robust optimization.

Time Slots \Γ	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
06/01/2017 00	62.2	60.9	57.1	55.8	49.5	49.2	39.9	30.7	24.5	20.6	17.2
06/01/2017 00	52.3	49.9	45.1	32.1	25.6	23.1	21.8	19.2	12.8	10.8	9.7
06/01/2017 17	327.3	324.3	320.5	307.9	301.1	294.5	282.3	264.1	240.6	218.6	190.8
06/01/2017 17	317.0	316.3	310.1	306.5	299.5	288.9	279.9	256.0	231.9	187.4	174.8
06/03/2017 00	182.5	179.3	178.1	176.1	175.9	172.7	160.1	153.8	135.3	119.8	103.8
06/03/2017 00	128.4	125.1	118.5	116.1	114.7	100.9	94.2	93.9	77.5	66.9	64.5
06/03/2017 17	280.1	279.9	276.8	259.2	249.7	243.5	234.0	206.9	194.4	185.3	157.1
06/03/2017 17	277.2	267.9	258.8	243.9	222.9	207.1	185.8	163.9	147.9	125.8	106.8
06/18/2017 00	148.7	147.9	146.6	142.5	142.1	140.9	139.1	136.9	129.9	121.1	105.5
06/18/2017 00	115.4	115.3	113.6	111.8	111.6	108.4	99.1	89.0	79.4	75.9	75.4
06/18/2017 17	286.4	280.2	276.7	273.4	269.9	269.5	257.4	254.0	241.4	237.9	224.8
06/18/2017 17	225.5	222.6	213.3	210.8	207.6	201.5	195.4	180.1	174.1	162.0	155.8
Avg.	214.5	195.4	209.3	202.5	198.0	195.1	185.5	174.4	161.0	150.6	133.2
Avg.	185.9	182.8	176.5	170.2	163.6	154.9	146.0	133.7	120.6	104.8	97.8



(a) June 1st (b) June 3rd (c) June 18th

Fig. 3: Comparison of violation rates by data-driven robust optimization and non-data-driven robust optimization

- We manually select the type of uncertainty set (i.e. polyhedral) for the robust optimization model. A smarter way is to let the data set to determine the appropriate type of uncertainty set. To this end, we will study a more intelligent strategy to select an appropriate uncertainty set for the given data set. In addition, the strategy of combining different types of uncertainty sets is also a potential research point.
- We study one-to-one driver and rider ride-sharing travel time savings for sake of simplicity, however, our proposed data-driven robust optimization model could be readily extended to multiple drivers and riders matching. We will investigate this research field in the future.

REFERENCES

[1] V. Armant and K. N. Brown, "Minimizing the driving distance in ride sharing systems," in *2014 IEEE 26th International Conference on Tools with Artificial Intelligence*. IEEE, 2014, pp. 568–575.

[2] Y. Duan, N. Wang, and J. Wu, "Optimizing order dispatch for ride-sharing systems," in *2019 28th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2019, pp. 1–9.

[3] Y. Li and S. H. Chung, "Ride-sharing under travel time uncertainty: Robust optimization and clustering approaches," *Computers & Industrial Engineering*, vol. 149, p. 106601, 2020.

[4] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton university press, 2009.

[5] Y. Guo, Y. Zhang, and Y. Boulaksil, "Real-time ride-sharing framework with dynamic timeframe and anticipation-based migration," *European Journal of Operational Research*, vol. 288, no. 3, pp. 810–828, 2021.

TABLE IV: The comparison of average travel time savings for the ride-sharing systems (in minutes) by data-driven and non-data-driven robust optimization under different levels of uncertainty degree. For each time slot, the top row is derived from data-driven robust optimization, the bottom row is derived from non-data-driven robust optimization, and the change compared to non-data-driven robust optimization is in the middle in italics.

Time Slots \Γ	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
06/01/2017 00	1.88	1.85	1.73	1.69	1.50	1.49	1.21	0.93	0.74	0.62	0.52
	(*16.8%*)	(*22.5%*)	(*27.2%*)	(*74.2%*)	(*92.3%*)	(*112.8%*)	(*83.3%*)	(*60.3%*)	(*89.7%*)	(*87.9%*)	(*79.3%*)
06/01/2017 00	1.61	1.51	1.36	0.97	0.78	0.7	0.66	0.58	0.39	0.33	0.29
06/01/2017 17	1.57	1.55	1.53	1.47	1.44	1.41	1.35	1.26	1.15	1.05	0.91
	(*3.3%*)	(*2.6%*)	(*3.8%*)	(*1.0%*)	(*1.0%*)	(*2.2%*)	(*1.5%*)	(*3.3%*)	(*3.6%*)	(*18.0%*)	(*8.3%*)
06/01/2017 17	1.52	1.51	1.48	1.46	1.43	1.38	1.33	1.22	1.11	0.89	0.84
06/03/2017 00	1.79	1.76	1.74	1.73	1.72	1.69	1.57	1.51	1.33	1.17	1.02
	(*42.1%*)	(*44.3%*)	(*20.7%*)	(*54.5%*)	(*53.6%*)	(*70.7%*)	(*70.6%*)	(*64.1%*)	(*75.0%*)	(*77.3%*)	(*61.9%*)
06/03/2017 00	1.26	1.22	1.16	1.14	1.12	0.99	0.92	0.92	0.76	0.66	0.63
06/03/2017 17	1.77	1.77	1.75	1.64	1.58	1.54	1.48	1.31	1.23	1.17	0.99
	(*1.1%*)	(*4.7%*)	(*6.7%*)	(*6.5%*)	(*12.1%*)	(*17.5%*)	(*25.4%*)	(*25.9%*)	(*30.8%*)	(*48.1%*)	(*47.8%*)
06/03/2017 17	1.75	1.69	1.64	1.54	1.41	1.31	1.18	1.04	0.94	0.79	0.67
06/18/2017 00	2.21	2.20	2.18	2.13	2.12	2.10	2.08	2.04	1.94	1.81	1.57
	(*28.5%*)	(*27.9%*)	(*29.0%*)	(*27.5%*)	(*26.9%*)	(*29.6%*)	(*40.5%*)	(*53.4%*)	(*63.0%*)	(*60.2%*)	(*40.2%*)
06/18/2017 00	1.72	1.72	1.69	1.67	1.67	1.62	1.48	1.33	1.19	1.13	1.12
06/18/2017 17	1.87	1.83	1.81	1.79	1.76	1.76	1.68	1.66	1.58	1.55	1.47
	(*27.2%*)	(*26.2%*)	(*30.2%*)	(*30.7%*)	(*30.4%*)	(*33.3%*)	(*31.3%*)	(*40.7%*)	(*38.6%*)	(*46.2%*)	(*44.1%*)
06/18/2017 17	1.47	1.45	1.39	1.37	1.35	1.32	1.28	1.18	1.14	1.06	1.02
Avg.	1.85	1.83	1.79	1.74	1.69	1.67	1.56	1.45	1.33	1.23	1.08
Avg.	(*19.4%*)	(*20.4%*)	(*23.4%*)	(*27.9%*)	(*31.0%*)	(*36.9%*)	(*36.8%*)	(*39.4%*)	(*44.6%*)	(*51.9%*)	(*42.1%*)
Avg.	1.55	1.52	1.45	1.36	1.29	1.22	1.14	1.04	0.92	0.81	0.76

[6] Q. Lin, L. Deng, J. Sun, and M. Chen, "Optimal demand-aware ride-sharing routing," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 2699–2707.

[7] X. Wang, N. Agatz, and A. Erera, "Stable matching for dynamic ride-sharing systems," *Transportation Science*, vol. 52, no. 4, pp. 850–867, 2018.

[8] Y. Wang, Y. Zhang, and J. Ma, "Dynamic real-time high-capacity ride-sharing model with subsequent information," *IET Intelligent Transport Systems*, vol. 14, no. 7, pp. 742–752, 2020.

[9] T.-Y. Ma, S. Rasulkhani, J. Y. Chow, and S. Klein, "A dynamic ridesharing dispatch and idle vehicle repositioning strategy with integrated transit transfers," *Transportation Research Part E: Logistics and Transportation Review*, vol. 128, pp. 417–442, 2019.

[10] B. Sun, M. Wei, and W. Wu, "An optimization model for demand-responsive feeder transit services based on ride-sharing car," *Information*, vol. 10, no. 12, p. 370, 2019.

[11] J. Farhan and T. D. Chen, "Impact of ridesharing on operational efficiency of shared autonomous electric vehicle fleet," *Transportation Research Part C: Emerging Technologies*, vol. 93, 2018.

[12] S. Jain and P. Biyani, "Improved real time ride sharing via graph coloring," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 956–961.

[13] A. Simonetto, J. Monteil, and C. Gambella, "Real-time city-scale ridesharing via linear assignment problems," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 208–232, 2019.

[14] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[15] A. Ben-Tal and A. Nemirovski, "Robust optimization–methodology and applications," *Mathematical programming*, vol. 92, no. 3, pp. 453–480, 2002.

[16] C. Ning and F. You, "Data-driven robust milp model for scheduling of multipurpose batch processes under uncertainty," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 6180–6185.

[17] G. E. Box and D. A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *Journal of the American statistical Association*, vol. 65, no. 332, pp. 1509–1526, 1970.

[18] X. Li, C. Wang, and X. Huang, "Reducing car-sharing relocation cost through non-parametric density estimation and stochastic programming," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–6.

[19] J. Gao, X. Li, C. Wang, and X. Huang, "Learning-based open driver guidance and rebalancing for reducing riders' wait time in ride-hailing platforms," in *2020 IEEE International Smart Cities Conference (ISC2)*. IEEE, pp. 1–7.

[20] F. Miao, S. He, L. Pepin, S. Han, A. Hendawi, M. E. Khalefa, J. A. Stankovic, and G. Pappas, "Data-driven distributionally robust optimization for vehicle balancing of mobility-on-demand systems," *ACM Transactions on Cyber-Physical Systems*, vol. 5, no. 2, pp. 1–27, 2021.

DATA-DRIVEN KALMAN-BASED VELOCITY ESTIMATION FOR AUTONOMOUS RACING

Adrià López Escoriza, Guy Revach, Nir Shlezinger, and Ruud J. G. van Sloun

ABSTRACT

Real-time velocity estimation is a core task in autonomous driving, which is carried out based on available raw sensors such as wheel odometry and motor currents. When the system dynamics and observations can be modeled together as a fully known linear Gaussian state space (SS) model, the celebrated Kalman filter (KF) is a low complexity optimal solution. However, both linearity of the underlying SS model and accurate knowledge of it are often not encountered in practice. This work proposes to estimate the velocity using a hybrid data-driven (DD) implementation of the KF for non-linear systems, coined KalmanNet. KalmanNet integrates a compact recurrent neural network in the flow of the classical KF, retaining low computational complexity, high data efficiency, and interpretability, while enabling operation in non-linear SS models with partial information. We apply KalmanNet on an autonomous racing car as part of the Formula Student (FS) Driverless competition. Our results demonstrate the ability of KalmanNet to outperform a state-of-the-art implementation of the KF that uses a postulated SS model, while being applicable on the vehicle control unit used by the car.

Index Terms— Kalman filter, autonomous vehicles.

1. INTRODUCTION

Many emerging technologies carry out real-time state estimation of dynamical systems. For instance, velocity estimation (VE) is a key and central component of autonomous vehicles [1]. The velocity estimates are used by the perception algorithms to compensate for motion when sensing the environment; they are required to provide localization and mapping; and are essential for providing feedback to the vehicle motion control system. VE must therefore provide high-rate, high-quality data. Due to the inertial nature of velocity, directly measuring it involves using expensive external velocity sensors such as optical flow sensors. While these sensors provide robustness against extreme conditions, their high cost limits their applicability for commercial road cars. It is therefore desired to design robust and accurate VE mechanisms that can process sensory data acquired by the vehicle, at a reduced cost, while operating reliably in extreme scenarios like adverse weather without requiring driver attention, thus reaching full (level 4) autonomy [2].

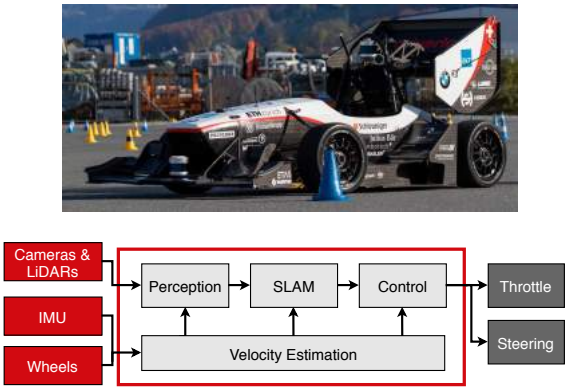


Fig. 1: Top: *pilatus* driverless, the Formula Student (FS) race car used. Bottom: Simplified software architecture.

The generic problem of state estimation is typically tackled using either of two leading strategies. The most widely used solution is based on the celebrated Kalman filter (KF) [3]. The KF is the minimum mean-squared error (MMSE) estimator in linear, time-invariant, state space (SS) models with Gaussian noise, and requires the model to be fully known. In many practical setups, including VE in self-driving cars, this model assumption may not hold; the underlying dynamics are complex and non-linear, while domain knowledge often relies on a crude approximation. Well-known model-based (MB) variants of the KF are designed for non-linear dynamics such as the extended Kalman filter (EKF) [4, Ch. 7] and the unscented Kalman filter (UKF) [5] but they are not MMSE optimal and are severely degraded by model mismatch [6]. An alternative strategy is to learn the filter mapping from data. In particular, neural network (NN) architectures such as recurrent neural networks (RNNs) have been shown to learn to carry out time series predication [7]. The work [8] trained an RNN outperforming MB KF architectures. While such data-driven (DD) architectures can learn to capture complex dynamics, they tend to require many trainable parameters even for seemingly simple sequence models [9], and may be computationally prohibitive to implement on limited hardware. In addition, these black-box architectures provide results that are not explainable enough, which makes them unfit for critical applications. These constraints limit the application of highly parameterized deep models for real-time state estimation in

A. L. Escoriza and G. Revach are with the Signal Processing Laboratory (ISI), Department of Information Technology and Electrical Engineering, ETH Zürich, Switzerland (e-mail: alopez@student.ethz.ch; grevach@ethz.ch). N. Shlezinger is with the School of ECE, Ben-Gurion University of the Negev, Beer Sheva, Israel (e-mail: nirshl@bgu.ac.il). R. J. G. van Sloun is with the EE Dpt., Eindhoven University of Technology, and with Phillips Research, Eindhoven, The Netherlands (e-mail: r.j.g.v.sloun@tue.nl). We thank the members of the AMZ Driverless team.

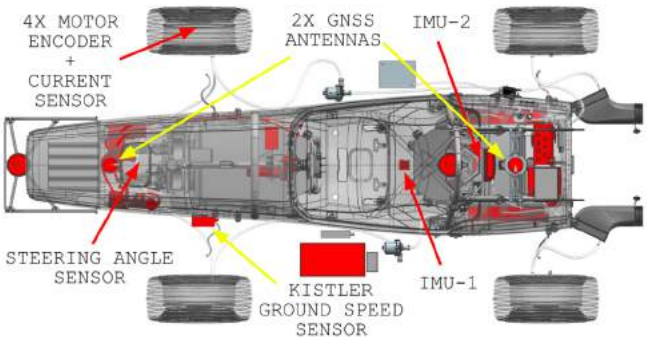


Fig. 2: Sensor setup: two IMUs, four motor encoders + current sensors, a steering angle sensor, and two dedicated velocity sensors that are used for validation and target generation.

applications embedded on hardware-limited mobile devices such as VE in autonomous vehicles.

In this work we study VE for self-driving cars using hybrid MB/DD state estimation [10]. Our design is based on the recent KalmanNet algorithm [11], which combines the soundness and low complexity of the classic KF while exploiting the model-agnostic nature of NNs to mitigate its dependence on accurate knowledge of the SS model. Our hybrid approach allows obtaining reliable velocity estimates in real time without relying on expensive velocity sensors. The method is demonstrated on a full scale autonomous race car, illustrated in Fig. 1, which can accelerate from 0 to 100 $[\mathrm{km/h}]$ in 2.1 s and reaches lateral accelerations of 1.7 g. KalmanNet operates at reduced complexity with a limited number of trainable parameters. It can thus be applied on the hardware-limited Vehicle Control Unit (VCU) of the autonomous car, as opposed to previously proposed purely DD estimators whose applicability was limited by their complexity [8]. We train the hybrid MB/DD KalmanNet to track the velocity of the car based solely on sensory inputs acquired from inertial measurement units (IMUs), wheel odometry, and motor currents; i.e., without relying on expensive velocity sensors. The resulting system is shown to notably outperform State-of-the-Art (SOA) KF with equivalent sensor setups, while approaching the accuracy achievable with costly sensory data. Our results demonstrate the potential of combining MB and DD methods in autonomous systems, while providing a proof of concept for KalmanNet in a challenging real-life scenario.

The rest of this paper is organized as follows: Section 2 discusses the system model, and Section 3 details the KalmanNet velocity estimator. Our experimental study is presented in Section 4, and Section 5 concludes the paper.

2. SYSTEM MODEL AND PRELIMINARIES

2.1. Velocity Estimation Problem Formulation

We consider the real-time VE problem in which at every time instance t our goal is to estimate a velocity vector $\mathbf{x}_t$ from a vector of sensory measurements $\mathbf{y}_t$ that are acquired by the autonomous vehicle. In particular,

$$\mathbf{x}_t = \left(\mathbf{v}_t^\top, \dot{\psi}_t, \mathbf{a}_t^\top\right)^\top \in \mathbb{R}^m, \quad m = 5 \tag{1}$$

where, $\mathbf{v}_t = (v_{x,t}, v_{y,t})^\top$ and $\mathbf{a}_t = (a_{x,t}, a_{y,t})^\top$ denotes the velocity and acceleration along the $x-y$ axis, respectively, and $\dot{\psi}_t$ denotes the rotational (yaw) rate along the z-axis. For the estimation task, we consider $\tilde{n} = 18$ raw measurements that are collected by sensors on the autonomous car. Two IMUs measure yaw rate and accelerations $\tilde{\mathbf{y}}_{\text{IMU}} = (\tilde{\dot{\psi}}, \tilde{a}_x, \tilde{a}_y)^\top$. Four *motor encoders* measure the rotational velocity of the i-th wheel ω_i. One *steering angle sensor* measures the steering angle of the wheels δ_i. One current sensor measures the torque T_i of the i-th wheel. The data from the wheel encoders is fused to get a velocity estimate of the i-th wheel. The car is also equipped with an optical flow-based velocity sensor and a GNSS velocity sensor, which are used only for validation and ground truth (GT). The sensor setup is depicted in Fig. 2.

2.2. State Space Model

To track the velocity, we assume that the evolution of the velocity state vector $\mathbf{x}_t$ together with the observed sensory data $\mathbf{y}_t$, can be described by a SS model. Specifically, as in [12], we assume that the evolution of the velocity is described by a set of continuous-time differential equations:

$$\dot{\mathbf{a}}_t = \mathbf{q}_{\mathbf{a},t} \qquad \ddot{\psi}_t = q_{\dot{\psi},t} \tag{2a}$$

$$\dot{v}_{x,t} = a_{x,t} + v_{y,t} \cdot \dot{\psi}_t \qquad \dot{v}_{y,t} = a_{y,t} - v_{x,t} \cdot \dot{\psi}_t \tag{2b}$$

where the change in the accelerations is modeled as an additive white Gaussian noise (AWGN) ($\mathbf{q}_\mathbf{a}$ and $q_{\dot{\psi}}$). Using standard techniques [13] we obtain the discrete-time evolution model:

$$\mathbf{x}_t = \mathbf{F}_t \cdot \mathbf{x}_{t-1} + \mathbf{q}_t, \qquad \mathbf{q}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}). \tag{3}$$

Here, $\mathbf{Q}$ is an unknown covariance matrix, and $\mathbf{F}_t$ is a time-dependent evolution matrix with $\Delta t = 5$ $[\mathrm{ms}]$

$$\mathbf{F}_t = \begin{pmatrix} 1 & \dot{\psi}_{t-1} \cdot \Delta t & 0 & \Delta t & 0 \\ -\dot{\psi}_{t-1} \cdot \Delta t & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

We assume a partially observable system, where the observations vector $\mathbf{y}_t$ is generated from $\mathbf{x}_t$ via

$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{n}_t, \quad \mathbf{n}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}). \tag{5}$$

Here, $\mathbf{R}$ is an unknown covariance matrix. The hidden states are observed by the sensors in the following way: accelerations and rotation rate are noisy measurements directly sensed

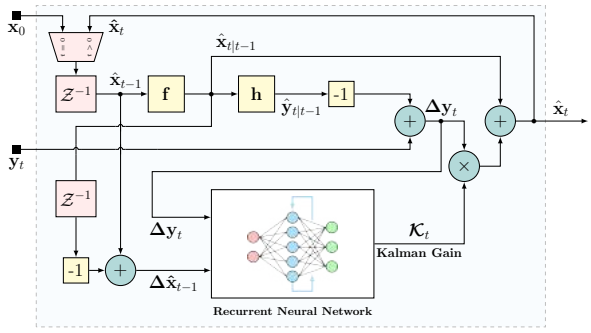


Fig. 3: KalmanNet block diagram.

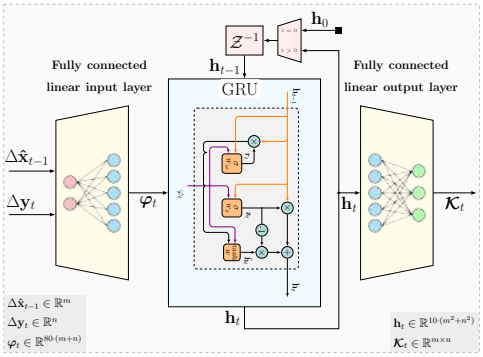


Fig. 4: KalmanNet neural network block diagram.

by the two IMUs

$$\left(\tilde{\dot{\psi}}_t, \tilde{a}_{x,t}, \tilde{a}_{y,t}\right)^\top = \mathbf{I}_{3\times 3} \cdot \left(\dot{\psi}_t, a_{x,t}, a_{y,t}\right)^\top + \mathbf{n}_{\text{IMU},t} \quad (6)$$

As in [14], measurements from the wheel encoders are fused to get an estimate of the wheel velocity

$$\hat{\mathbf{v}}_{i,t} = \left[\cos\left(\delta_{i,t}\right), \sin\left(\delta_{i,t}\right)\right]^\top \cdot \omega_{i,t} \cdot \frac{R_i}{\text{SR}\left(T_{i,t}\right) + 1}. \quad (7)$$

Here, $\text{SR}(\cdot)$ is a function that maps the torque to its slip ratio under low slip conditions (Pacejka magic tire model [15]), and R_i is the radius of the i-*th* wheel. A noisy observation for the car velocity is now assumed to be given by

$$i_t^* = \arg\min_{i\in\{1,..,4\}} \{\hat{v}_{i,x,t}\} \quad (8a)$$

$$\hat{\mathbf{v}}_{i_t^*,t} = \mathbf{v}_t + \left[\dot{\psi}_{t-1} \cdot p_{i_t^*,y}, -\dot{\psi}_{t-1} \cdot p_{i_t^*,x}\right]^\top + \mathbf{n}_{\mathbf{v},t}. \quad (8b)$$

Here, $\mathbf{n}_{\mathbf{v},t}$ is AWGN, $p_{i,x}$ and $p_{i,y}$ are constants denoting the positions of the i-th wheel in the car frame, and i_t^* is the index of the wheel with the lowest estimated x-axis velocity; i.e., slip. The reason that the velocity update only takes the single wheel with the smallest absolute estimated slip stems from the fact that the slip ratio calculation is fairly accurate at low slips, but uncertain at high slips. By using this model we can use $n = 8$ fused observation as input to the filter.

2.3. Kalman Filtering in Velocity Estimation

We compare the performance of KalmanNet to a Mixed Kalman Filter (MKF) [8]. It is a MB approach that uses a combination of EKF, UKF, and a chi-squared test for detecting outlier measurements. MKF is the filter that was embedded in the autonomous car (*pilatus*), it is well tuned and optimized, it was extensively tested and proved to be successful in multiple FS competitions around Europe in 2019 and during several testing sessions in different kind of tracks in 2020, and is therefore considered SOA for VE in autonomous racing cars. In the MKF the state is first propagated with an EKF step, then a chi-squared test is done on the measurements. The update step is performed via EKF by default except for the wheelspeed measurements, which, due to the strong non-linearity of the model, are updated with a UKF step.

3. HYBRID MB/DD VELOCITY ESTIMATION

Our *hybrid* velocity estimator is based upon our KalmanNet architecture presented in detail in [11]. We first briefly describe KalmanNet, then explain how it was applied to VE and discuss its properties. Finally we discuss the gains of KalmanNet compared to related architectures on a conceptual level.

3.1. KalmanNet

KalmanNet is built upon the flow of the KF. However, as opposed to classic KF that utilizes full domain knowledge to compute the Kalman gain (KG), KalmanNet learns the KG from data by training a compact RNN in an end-to-end (E2E) manner using GT state vectors. This adaptability to the data provides KalmanNet with robustness to model mismatch and removes the need to know the noise covariance matrices. A block diagram of KalmanNet, where the learned KG is integrated in the overall KF flow, is illustrated in Fig. 3. Compared with E2E black-box architectures, KalmanNet inherits the interpretability from the KF, providing explainable results, by learning how to perform Kalman Filtering rather than directly learning the dependencies between states and observations. These capabilities of KalmanNet were extensively evaluated in [11] for linear system models and in [16] for nonlinear system models.

3.2. KalmanNet Velocity Estimator

We adapt KalmanNet architecture to the task of VE by integrating the postulated SS model detailed in Section 2.2. Note that this integration involves only setting the mapping blocks $\mathbf{f}$ and $\mathbf{h}$, as knowledge of the noise statistics is not required by KalmanNet, which learns to compute the KG directly from data. We include a prepossessing block that performs sensor calibration and frame transformations based on the design specifications of the autonomous car and ignores the first time steps of the sensors to use them for bias calibration. We use a single gated recurrent unit cell with a hidden state of size proportional to the size of $\mathbf{Q}$ and $\mathbf{R}$. This reduced architecture allows the system to run in the VCU of the car at the high

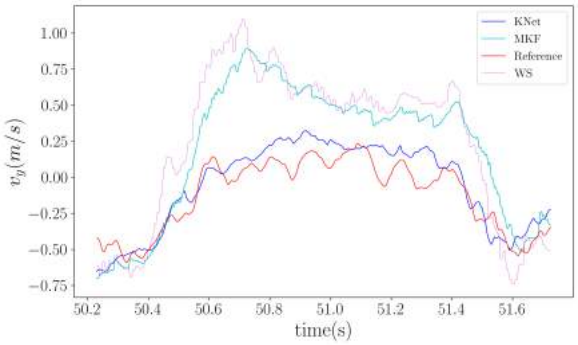


Fig. 5: Tracking v_y compared to sensor inputs and MKF.

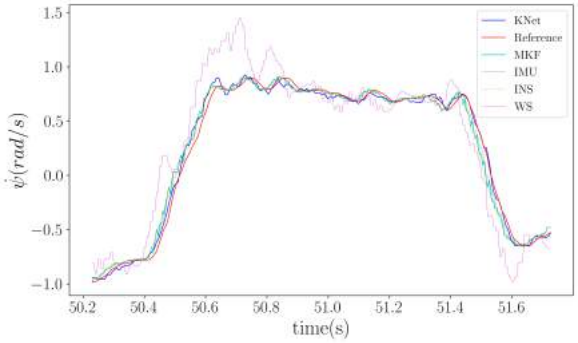


Fig. 6: Tracking $\dot{\psi}$ compared to sensor inputs and MKF.

frequency required by the estimator (around 200 [Hz]). An illustration of the RNN architecture is shown in Fig. 4.

3.3. Discussion

While classical KF architectures are proven robust and stable solutions for most estimation problems, they are limited by the model mismatch inherent to most real world applications. This model mismatch usually imposes an error floor on the accuracy of the estimator making them sub-optimal for challenging applications like autonomous driving. KalmanNet can overcome this error floor by adapting to the application through data, learning how to obtain the optimal performance of the KF. Secondly, as mentioned in Section 1, the pivotal role of VE in the autonomous system imposes a requirement for robustness and stability. While E2E DD architectures can accurately learn the complex dynamics of the system, the lack of explainability in their results makes them a risky approach for such critical applications. Taking the best of both worlds, KalmanNet fuses the interpretability of the classical KF with the adaptability of RNN to provide the autonomous system with accurate and reliable estimations for velocity.

4. EXPERIMENTS AND RESULTS

All the experiments were carried out using real data collected by the autonomous race car *pilatus*, developed by the members of AMZ Driverless across three consecutive FS seasons. Datasets were created from real data obtained from noisy sensors over both testing and competition runs on FS style tracks. The raw sensor measurements have different sampling frequencies of 200 [Hz] (IMU1 and steering angle), 125 [Hz] (IMU2), and 100 [Hz] (wheel speeds and torques). These raw measurements are hardware time-synced by sampling at a constant rate of 200 [Hz] (zero-order hold) and used as input to KalmanNet. Following the approach used in [8], we generate the reference for KalmanNet using the MKF detailed in Section 2.3 in combination with the costly external velocity sensors described in Section 2.1. The target of the KalmanNet approach should thus obtain similar results but do so without relying on expensive velocity sensors. The output from the MKF was also post-processed using a non-causal, Gaussian moving average filter to obtain a smoothed, non-delayed target, referred to as the reference. As a benchmark, we consider the MKF operating with the same sensors as KalmanNet; i.e., without the external velocity sensors. The mean-squared error of KalmanNet computed over the entire dataset compared to MKF is summarized in Table 1, while Figs. 5-6 show the tracking of the lateral velocity v_y and the yaw rate $\dot{\psi}$ for data collected in a test run in Tuggen (Switzerland) in August 2020. We observe in Table 1 that KalmanNet outperforms the baseline in 4 of the 5 trackable states. Of particular interest are the gains in tracking v_y, since, due to the model inaccuracy at high slips, this is typically the most challenging variable to trace. Fig. 5 shows how while the baseline gets mislead by the inaccurate wheel-speed (WS) measurements, KalmanNet obtains a smoother estimation much closer to the reference; and the numerical evaluation over the entire dataset shows a 3.63 [dB] gain over the baseline. Fig. 6 shows that KalmanNet also manages to get a more accurate trajectory when we have accurate sensors, achieving a 4.42 [dB] gain over the baseline.

Table 1: KalmanNet performance compared to MKF and raw measurements in terms of RMSE normalized to each state.

State	MKF		KalmanNet		Obs	
	linear	dB	linear	dB	linear	dB
v_x	0.32	**-69.88**	0.68	-63.24	0.68	-63.24
v_y	9.26	-40.66	6.10	**-44.29**	10.93	-39.22
$\dot{\psi}$	1.13	-58.96	0.60	**-64.38**	1.72	-55.28
a_x	2.57	-51.81	2.41	**-52.36**	4.55	-46.84
a_y	2.04	-53.80	1.67	**-55.56**	3.62	-48.82

5. CONCLUSIONS

In this work we studied VE in autonomous racing cars using KalmanNet architecture. As KalmanNet uses a relatively compact NN it is applicable to computationally limited devices such as the VCU of most cars. Our experimental results show that KalmanNet approaches the performance of the SOA MKF with additional external velocity sensors, and outperforms it when working with the same sensors.

6. REFERENCES

[1] Juraj Kabzan et al., "AMZ driverless: The full autonomous racing system," *Journal of Field Robotics*, vol. 37, no. 7, pp. 1267–1294, 2020.

[2] On-Road Automated Driving (ORAD) committee, *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Sept. 2016.

[3] Rudolph Emil Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[4] Simon S Haykin, *Adaptive Filter Theory*, Pearson Education India, 2005.

[5] Simon J Julier and Jeffrey K Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*. International Society for Optics and Photonics, 1997, vol. 3068, pp. 182–193.

[6] Eric A Wan and Rudolph Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*. IEEE, 2000, pp. 153–158.

[7] Hojjat Salehinejad, Sharan Sankar, Joseph Barfett, Errol Colak, and Shahrokh Valaee, "Recent advances in recurrent neural networks," 2018.

[8] Sirish Srinivasan, Inkyu Sa, Alex Zyner, Victor Reijgwart, Miguel I Valls, and Roland Siegwart, "End-to-end velocity estimation for autonomous racing," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6869–6875, 2020.

[9] Manzil Zaheer, Amr Ahmed, and Alexander J Smola, "Latent LSTM allocation: Joint clustering and non-linear dynamic modeling of sequence data," in *International Conference on Machine Learning*, 2017, pp. 3967–3976.

[10] Nir Shlezinger, Jay Whang, Yonina C Eldar, and Alexandros G Dimakis, "Model-based deep learning," *arXiv preprint arXiv:2012.08405*, 2020.

[11] Guy Revach, Nir Shlezinger, Ruud J. G. van Sloun, and Yonina C. Eldar, "KalmanNet: Data-driven Kalman filtering," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 3905–3909.

[12] RK Douglas, WH Chung, DP Malladi, RH Chen, DL Mingori, and JL Speyer, *Fault Detection and Identification with Application to Advanced Vehicle Control Systems. Final Report*, chapter 8, 1996.

[13] Yaakov Bar-Shalom, X Rong Li, and Thiagalingam Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*, John Wiley & Sons, 2004.

[14] L. Imsland, T.A. Johansen, T.I. Fossen, J.C. Kalkkuhl, and A. Suissa, "Vehicle velocity estimation using modular nonlinear observers," in *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005, pp. 6728–6733.

[15] H. Pacejka and I.J.M. Besselink, *Tire and Vehicle Dynamics*, Elsevier Science, 2012.

[16] Guy Revach, Adrià López Escoriza, Nir Shlezinger, Ruud J G van Sloun, and Yonina C Eldar, *KalmanNet: Neural Network Aided Kalman Filtering for Non-Linear Dynamics with Partial Domain Knowledge*, 2021, *preprint* `http://people.ee.ethz.ch/~grevach/KalmanNet_NL_v1.pdf`.

Cooperative UWB-Based Localization for Outdoors Positioning and Navigation of UAVs aided by Ground Robots

Yu Xianjia†, Li Qingqing†, Jorge Peña Queralta†, Jukka Heikkonen†, Tomi Westerlund†

†Turku Intelligent Embedded and Robotic Systems (TIERS) Lab, University of Turku, Finland.
Emails: [1]{xianjia.yu, qingqli, jopequ, jukhei, tovewe}@utu.fi

***Abstract*—Unmanned aerial vehicles (UAVs) are becoming largely ubiquitous with an increasing demand for aerial data. Accurate navigation and localization, required for precise data collection in many industrial applications, often relies on RTK GNSS. These systems, able of centimeter-level accuracy, require a setup and calibration process and are relatively expensive. This paper addresses the problem of accurate positioning and navigation of UAVs through cooperative localization. Inexpensive ultra-wideband (UWB) transceivers installed on both the UAV and a support ground robot enable centimeter-level relative positioning. With fast deployment and wide setup flexibility, the proposed system is able to accommodate different environments and can also be utilized in GNSS-denied environments. Through extensive simulations and test fields, we evaluate the accuracy of the system and compare it to GNSS in urban environments where multipath transmission degrades accuracy. For completeness, we include visual-inertial odometry in the experiments and compare the performance with the UWB-based cooperative localization.**

***Index Terms*—UAV; GNSS; Ultra-wideband; UWB; VIO; Localization; MAV; UGV; Cooperative localization; Navigation;**

I. Introduction

Multiple industrial use cases benefit from the deployment of unmanned aerial vehicles (UAVs) [1]. When accurate localization is needed, GNSS-RTK is the de-facto standard for gathering aerial data with UAVs [2]. For example, high-accuracy photogrammetry [3], civil infrastructure monitoring [4], or in urban environments where GNSS signals suffer more degradation [2]. As UAVs become ubiquitous across different domains and application areas [5], having access to more flexible and lower-cost solutions to precise UAV navigation can aid in accelerating adoption and widespread use. In this paper, we consider the problem of UAV navigation through relative localization to a companion unmanned ground vehicle (UGV). We consider a ground robot as a more flexible platform from the point of view of deployment, but in simulations, we also consider localization based on fixed beacons in the environment, closer to how GNSS-RTK systems are deployed.

Within the different approaches that can be used for cooperative relative localization, from visual sensors [6] to cooperative SLAM [7], wireless ranging technologies offer high performance with low system complexity [8]. In particular, ultra-wideband (UWB) wireless ranging offers unparalleled localization performance within the different radio technologies in unlicensed bands [9]. Other benefits of UWB include resilience to multipath, high time resolution, and low interference with other radio technologies [10].

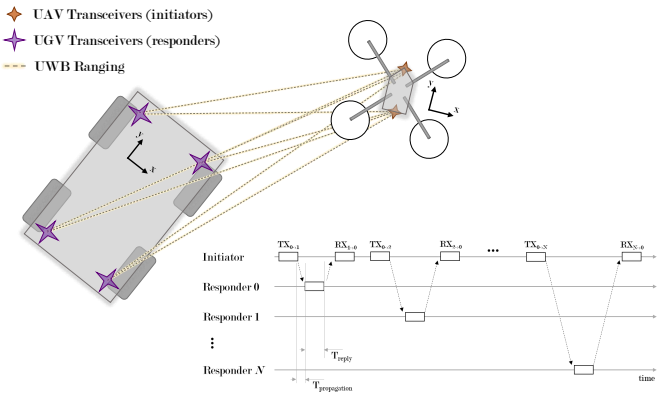


Fig. 1: Cooperative localization approach based on UWB ranging measurements from multiple transceivers in different robots

The system we analyze in this paper consists of a UGV equipped with four UWB transceivers and a UAV equipped with two transceivers. The UAV transceivers act as initiators, taking turns in sending signals to each of the UGV transceivers. When these respond, the time of flight of the signal is calculated and the distance between each pair of transceivers is calculated. This process is illustrated in Fig. 1. The main contribution of this paper is thus on evaluating how UWB-based relative localization can improve the positioning of UAVs when supported by ground robots. We simulate different trajectories to evaluate the performance of the system and compare the accuracy of the GNSS, UWB, and VIO approach to localization with field tests in an urban environment. In the simulations, we consider different configurations of transceivers in the ground to compare the localization and navigation performance.

The remainder of this document is organized as follows. Section II introduces absolute and relative positioning approaches relevant to the presented approach. Section III then describes the cooperative localization approach. In Section IV we introduce the methodology for simulations and ex-

periments, with results presented in Section V. Section VI concludes the work and outlines future research directions.

II. Background

This section reviews the literature in the area of outdoors positioning and navigation methods for multi-robot systems.

A. Limitations of standalone GNSS

The long-term operation of autonomous robots outdoors is often a reliance on GNSS [11]. However, the positioning accuracy of GNSS can be easily influenced by multipath when satellites are not in line-of-sight. This is a typical problem in urban environments or partly covered environments such as forests [12], [13]. Additional sensors are thus used in practice, from IMUs at the lowest level [14] to odometry estimation from lidars [15] or visual sensors [16]. It is worth mentioning, nonetheless, that more recent receivers exploiting multi-constellation signals (e.g., GPS, GLONASS, BEIDOU, or GALILEO) are able to deliver significantly higher positioning accuracy [17].

B. GNSS-RTK

High-accuracy GNSS positioning is possible with real-time kinematic (RTK) systems. RTK positioning leverages measurements of the phase of the signal's carrier wave in addition to the information content of the signal and relies on a reference station or interpolated virtual station to provide real-time corrections, providing up to centimeter-level accuracy [18]. These systems, however, are costly and require calibration and setup for each different location.

C. Onboard navigation

With the increasing adoption of UAVs in recent years, the maturity of onboard estate estimation and localization has reached a point where it is standard in commercial systems. Onboard odometry and positioning are typically based on monocular or stereo vision (e.g., VINS-mono [19], vins-fusion [20]), but lidars are also effective in larger UAVs [21]. Passive visual sensors, however, have evident limitations in terms of environmental conditions (e.g., night operation) and in situations where there is a lack of features [22], [23]

D. UWB Localization

Ultra-wideband (UWB) positioning systems are being increasingly adopted for autonomous systems [10].UWB positioning systems based on a series of fixed nodes in known locations (or anchors), and ranging measurements between these and mobile nodes (or tags), can be used for consistent, long-term localization of mobile robots [24], [8].Compared to RTK-based localization systems, UWB systems can be utilized both indoors and outdoors, can be automatically calibrated [25] for ad-hoc deployment, and offer similar accuracies at much lower prices. UWB sensors also have a small form factor and are generally considered more energy efficient than other wireless solutions. Finally, UWB ranging is often combined with other sensors to add orientation estimation and increase the overall localization performance. Different approaches in the literature include fusion of UWB with IMU [26], VIO estimators [27], GPS [28], or lidar [29].

E. UWB for relative estimation

UWB ranging has been widely used for relative localization within multi-robot systems. For instance, in [30], the authors demonstrate a system where relative positioning between and UAV and a UGV is designed based on UWB transceivers installed on both robots. In subsequent works [27], a similar system is employed during docking maneuvers. Combined with vision sensors for the final docking, the autonomous approach of the UAV to the UGV relied on UWB ranging between transceivers in both robots. Relative localization between UAVs and UGVs has also been shown within the context of collaborative dense scene reconstruction [31]. In multi-UAV systems and UAV swarms, UWB ranging has been leveraged for swarm-level decentralized estate estimation [32], [33].

In general, terms, while UWB systems including those for relative localization have been widely studied in the literature, we see a lack of studies that quantitatively analyze how UWB-based relative estate estimation can improve GNSS positioning and navigation outdoors.

III. Cooperative UWB-Based Localization

We consider the problem of relative localization between a UAV and a UGV based on UWB ranging between transceivers installed onboard both robots. The objective is to leverage this relative localization to improve the accuracy of the UAV navigation outdoors. We are especially interested in improving the navigation performance in urban areas where the accuracy of GNSS sensors is degraded due to the signal being reflected at or occluded by nearby buildings.

Let us denote by I the set of N transceivers onboard the UAV. These will act as initiators, i.e., will actively transmit messages to initiate ranging measurements between them and the responder transceivers on the ground. We denote the latter ones by the set R of size M. An initial approach, which we implement, is to iteratively range between each initiator and the set of responders. If the number of nodes increases significantly, more scalable approaches can be used where, for example, a single initiator message is answered by several or all responders with different delays [34].

We model the UWB ranges between an initiator i and a responder j with

$$\mathbf{z}_{(i,j)}^{UWB} = \|\mathbf{p}_i(t) - \mathbf{q}_j(t)\| + \mathcal{N}(0,\, \sigma_{UWB}) \tag{1}$$

where $\mathbf{p}_i$ and $\mathbf{q}_j$ represent the positions of the initiator and responder transceivers, respectively, and $\mathcal{N}$ is Gaussian noise. Based on the ranges, different approaches to localization include, e.g., multilateration or a least squares estimator (LE). We implement the latter, and hence the position of each tag can be calculated based on the known anchor positions by

$$\mathbf{p}_i = \underset{\mathbf{p}\in\mathbb{R}^3}{\operatorname{argmin}} \sum_{j=0}^{M} \left(\mathbf{z}_{(i,j)}^{UWB} - \|\mathbf{p} - \mathbf{q}_j\|\right)^2 \tag{2}$$

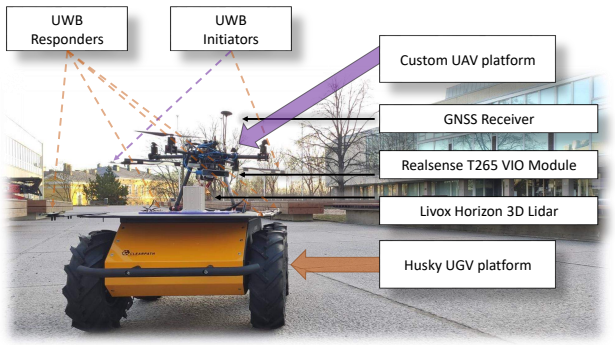


Fig. 2: UAV and companion ground robot utilized in the experiments.

Alternatively, assuming that the position of initiators in the UAV ($\{\mathbf{p}_i\}$) is given based on the UAV's position and orientation ($\mathbf{p}$ and θ, respectively) by a set of rigid body transformations f_i, i.e., $\mathbf{p}_i = f_i(\mathbf{p}, \theta)$, then the estimator can be used to obtain the full pose of the UAV directly with

$$\mathbf{p},\, \theta = \underset{\substack{\mathbf{p}\in\mathbb{R}^3 \\ \theta\in(-\pi,\pi]}}{\operatorname{argmin}} \sum_{i=0}^{N}\sum_{j=0}^{M} \left(\mathbf{z}_{(i,j)}^{UWB} - \|f_i(\mathbf{p},\theta) - \mathbf{q}_j\|\right)^2 \quad (3)$$

IV. Methodology

A. Simulation environment

The first tests are carried out in a simulation environment using ROS and Gazebo. We simulate the UWB ranging with a standard deviation of the Gaussian noise set to $\sigma_{UWB} = 10\,cm$. This is a conservative value based on the literature [8]. We simulate a single transceiver on the UAV and four transceivers on the ground. The latter ones are set are variable distances simulating deployment in small UGVs (0.6 separation), large UGVs (1.2 m separation) and different settings based, e.g., on tripods (with separations at 3 m, 4 m, 12 m and 16 m).

In the simulation experiments, we perform two types of flights. First, a vertical flight where the UAV is set to follow a straight vertical line up to an altitude of 30 m. Second, a flight following a square pattern with a fixed size of 8 by 8 m but at different altitudes (5 m, 10 m and 20 m). For each of these flights, we evaluate the UWB positioning performance with flights based on ground truth positioning. Then, we perform the flight using the UWB position estimation as control input and evaluate how well the UAV follows the predefined trajectory (we refer to this as navigation error).

B. Multi-robot system

The multi-robot system employed consists of a single ground robot and a UAV. The ground robot is a ClearPath Husky outdoor platform equipped with four UWB responder transceivers for cooperative positioning and a Livox Avia lidar utilized to obtain ground truth. Owing to the lack of a reference system such as a GNSS-RTK receiver, we extract the UAV position from the lidar's point cloud and utilize this as a reference. The point cloud is automatically processed following the steps described in Algorithm **??**, and manually validated. We refer the reader to [35]for further details on this method. Based on indoor testing with a reference anchor-based UWB system, we have evaluated the ground truth accuracy to be in the order of 10 cm. The UGV and the custom UAV are shown in Fig. 2. The UAV is equipped with two UWB transceivers and an Intel RealSense tracking camera T265 that performs VIO estimation.

C. Experimental settings

The field experiments are carried out in Turku, Finland (precise location is 60.4557389° N, 22.2843384° E), between a short line of trees and a large building that presumably blocks and reflects GNSS signals. The UAV runs the PX4 autopilot firmware, which is unable to obtain a stable GNSS lock in the test location. This location is chosen as an example of an urban location where GNSS receivers operate in suboptimal mode.

V. Experimental Results

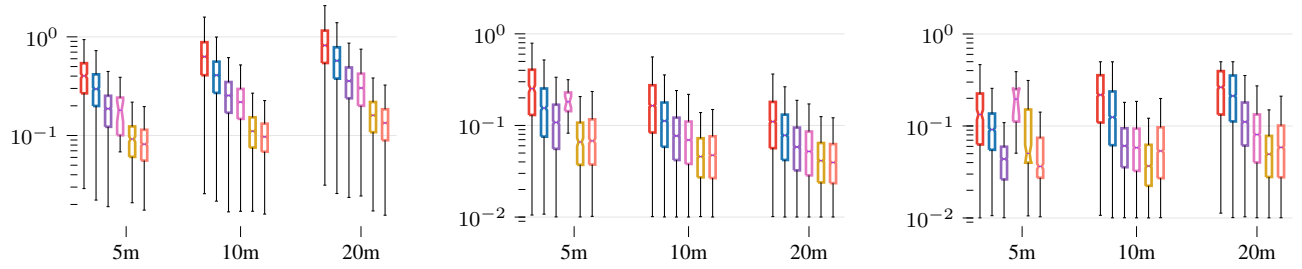
A. Simulation Results

The positioning and navigation errors for vertical flights are shown in Fig. 4. We observe that the positioning error consistently decreases as the anchors become more separated. For the small UGV setting, the error goes over 1 m almost 20% of the time, being highly unstable. It is worth noticing that the navigation error becomes relatively stable with the large UGV anchor distribution (1.2 m separation). Navigation errors are in general lower than their positioning counterparts as the control of the drone is less affected by individual ranging errors, and these tend to average to zero as time passes. It is also worth noticing that the altitude error is significantly lower in all cases when compared to the planar xy error.

Figure 3 then shows the results of flights following a square pattern. We can see that if UWB systems based on fixed anchors separated more than 10 m are utilized, then the navigation error can be consistently maintained below 10 cm. In the case of relying on small or large UGVs, the error is in the tens of centimeters, providing a competitive alternative to RTK-GNSS systems with higher deployment flexibility and lower system complexity.

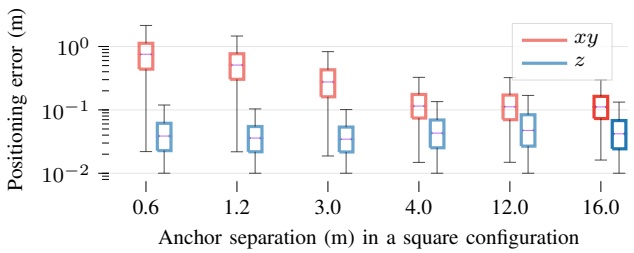
B. Experimental results

Results from outdoors experiments with real robots are reported in Fig. 5 and Fig. 6. The former shows a partial extract from the trajectory in 3D, where we can observe that the UWB error is significantly smaller even when the altitude reaches 30 m. In the latter plot we can see that the overall error more than 5 min flight time. The cooperative UWB approach particularly outperforms both VIO and GNSS estimations in terms of vertical accuracy. In terms of planar xy error, VIO is more accurate but only during the first few seconds of flight, before it rapidly loses accuracy and diverges when the UAV altitude increases. In any case, the cooperative UWB-based localization provides consistent accuracy throughout the flight and therefore has potential for better collection of aerial data through autonomous flights.

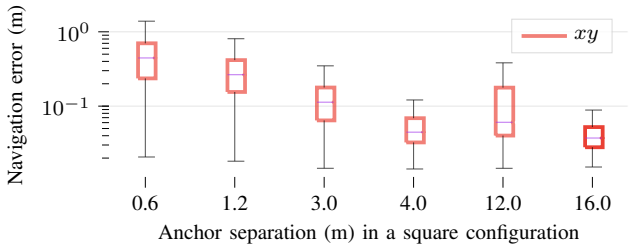


(a) Positioning xy error. (b) Positioning z error. (c) Navigation xy error.

Fig. 3: Positioning and navigation errors over a flight following a squared shape of 8 by 8 m, at three different altitudes (5, 10 and 20 m). The altitude is set to a constant so only the XY error is calculated for the UWB-based navigation. The legend has been omitted due to limited space, with the colors representing, from left to right in each group, anchors separated by 0.6 m, 1.2 m, 3 m, 4 m, 12 m and 16 m.



(a) Positioning error based on different anchor distribution settings when doing up and down navigation during a vertical flight.



(b) Navigation error based on different anchor distribution settings when doing up and down navigation during a vertical flight.

Fig. 4: Positioning and navigation errors over a vertical flight to an altitude of 30 m. The navigation error includes only the planar distance to the vertical line the drone is set to follow.

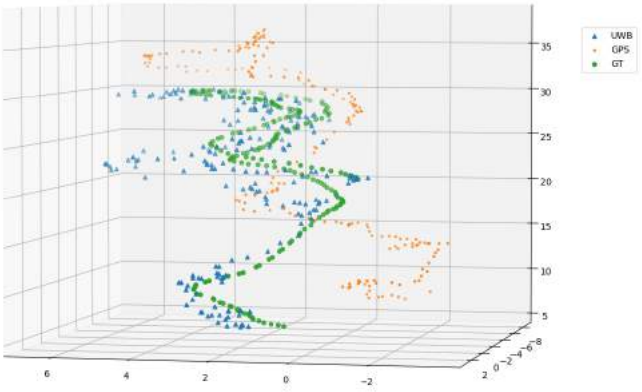


Fig. 5: Partial trajectory of the UAV during the outdoors experiment. VIO is not included because it becomes unusable once the UAV reaches 8 m of altitude.

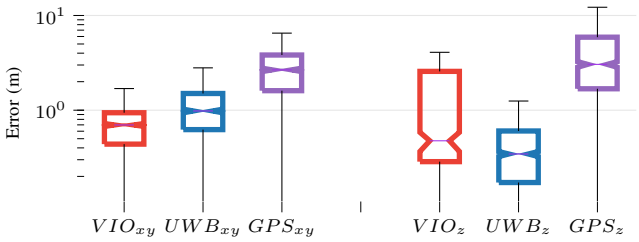


Fig. 6: Planar and vertical errors of the different methods during the outdoors flight. The VIO has low error but was only valid for the first few seconds of flight.

VI. Conclusion

We have presented an analysis on how UWB-based relative localization between a UAV and a companion ground robot can improve the accuracy of autonomous flights outdoors. In particular, we have simulated different scenarios to assess the accuracy of the UWB-based relative positioning method. We have then validated this with robots in outdoor experiments, in an urban area where GNSS receivers do not perform optimally. Our analysis includes VIO estimation, which is more accurate at first but loses the reference when the UAV starts gaining altitude, presumably due to the lack of reference points.

In summary, we can conclude that UWB-based positioning systems can provide an alternative to RTK-GNSS when the accuracy of standalone GNSS is not enough for gathering aerial data. Moreover, we have proved that even when the transceivers are placed near each other in the ground, mounted on a mobile platform, the accuracy is enough to enable autonomous flight.

Future work will focus on integrating UWB with GNSS data and performing experiments in more varied environments. We will also analyze different anchor configurations and consider additional robots.

Acknowledgment

This research work is supported by the Academy of Finland's AutoSOS project (Grant No. 328755) and RoboMesh project (Grant No. 336061).

REFERENCES

[1] Hazim Shakhatreh, Ahmad H Sawalmeh, Ala Al-Fuqaha, Zuochao Dou, Eyad Almaita, Issa Khalil, Noor Shamsiah Othman, Abdallah Khreishah, and Mohsen Guizani. Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges. *Ieee Access*, 7:48572–48634, 2019.

[2] Tuan Li, Hongping Zhang, Zhouzheng Gao, Qijin Chen, and Xiaoji Niu. High-accuracy positioning in urban environments using single-frequency multi-gnss rtk/mems-imu integration. *Remote sensing*, 10(2):205, 2018.

[3] Jae-One Lee and Sang-Min Sung. Assessment of positioning accuracy of uav photogrammetry based on rtk-gps. *Journal of the Korea Academia-Industrial cooperation Society*, 19(4):63–68, 2018.

[4] Kiyoung Kim, Jaemook Choi, Junyeon Chung, Gunhee Koo, In-Hwan Bae, and Hoon Sohn. Structural displacement estimation through multi-rate fusion of accelerometer and rtk-gps displacement and velocity measurements. *Measurement*, 130:223–235, 2018.

[5] Jorge Peña Queralta, Jussi Taipalmaa, Bilge Can Pullinen, Victor Kathan Sarker, Tuan Nguyen Gia, Hannu Tenhunen, Moncef Gabbouj, Jenni Raitoharju, and Tomi Westerlund. Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. *IEEE Access*, 8:191617–191643, 2020.

[6] Cheng Hui, Chen Yousheng, Li Xiaokun, and Wong Wing Shing. Autonomous takeoff, tracking and landing of a uav on a moving ugv using onboard monocular vision. In *Proceedings of the 32nd Chinese control conference*, pages 5895–5901. IEEE, 2013.

[7] Pileun Kim, Leon C Price, Jisoo Park, and Yong K Cho. Uav-ugv cooperative 3d environmental mapping. In *Computing in Civil Engineering 2019: Data, Sensing, and Analytics*, pages 384–392. American Society of Civil Engineers Reston, VA, 2019.

[8] Jorge Peña Queralta, Carmen Martínez Almansa, Fabrizio Schiano, Dario Floreano, and Tomi Westerlund. Uwb-based system for uav localization in gnss-denied environments: Characterization and dataset. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4521–4528, 2020.

[9] Wang Shule, Carmen Martínez Almansa, Jorge Peña Queralta, Zhuo Zou, and Tomi Westerlund. Uwb-based localization for multi-uav systems and collaborative heterogeneous multi-robot systems. *Procedia Computer Science*, 175:357–364, 2020.

[10] Xianjia Yu, Qingqing Li, Jorge Peña Queralta, and Tomi Westerlund. Applications of uwb networks and positioning to autonomous robots and industrial systems. *arXiv preprint arXiv:2103.13488*, 2021.

[11] Li Qingqing, Jorge Peña Queralta, Tuan Nguyen Gia, Zhuo Zou, and Tomi Westerlund. Multi sensor fusion for navigation and mapping in autonomous vehicles: Accurate localization in urban environments. *The 9th IEEE CIS-RAM*, 2019.

[12] Paul D Groves, Ziyi Jiang, Lei Wang, and Marek K Ziebart. Intelligent urban positioning using multi-constellation gnss with 3d mapping and nlos signal detection. In *Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)*, pages 458–472, 2012.

[13] Qingqing Li, Paavo Nevalainen, Jorge Peña Queralta, Jukka Heikkonen, and Tomi Westerlund. Localization in unstructured environments: Towards autonomous robots in forests with delaunay triangulation. *Remote Sensing*, 12(11):1870, 2020.

[14] San Jiang and Wanshou Jiang. On-board gnss/imu assisted feature extraction and matching for oblique uav images. *Remote Sensing*, 9(8):813, 2017.

[15] Le Chang, Xiaoji Niu, Tianyi Liu, Jian Tang, and Chuang Qian. Gnss/ins/lidar-slam integrated navigation system based on graph optimization. *Remote Sensing*, 11(9):1009, 2019.

[16] Tuan Li, Hongping Zhang, Zhouzheng Gao, Xiaoji Niu, and Naser El-Sheimy. Tight fusion of a monocular camera, mems-imu, and single-frequency multi-gnss rtk for precise navigation in gnss-challenged environments. *Remote Sensing*, 11(6):610, 2019.

[17] Xingxing Li, Xin Li, Gege Liu, Guolong Feng, Yongqiang Yuan, Keke Zhang, and Xiaodong Ren. Triple-frequency ppp ambiguity resolution with multi-constellation gnss: Bds and galileo. *Journal of Geodesy*, 93(8):1105–1122, 2019.

[18] Julián Tomaštík, Martin Mokroš, Peter Surovỳ, Alžbeta Grznárová, and Ján Merganič. Uav rtk/ppk method—an optimal solution for mapping inaccessible forested areas? *Remote sensing*, 11(6):721, 2019.

[19] Tong Qin, Peiliang Li, and Shaojie Shen. Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4):1004–1020, 2018.

[20] Tong Qin, Jie Pan, Shaozu Cao, and Shaojie Shen. A general optimization-based framework for local odometry estimation with multiple sensors. *arXiv preprint arXiv:1901.03638*, 2019.

[21] Ji Zhang, Chen Hu, Rushat Gupta Chadha, and Sanjiv Singh. Maximum likelihood path planning for fast aerial maneuvers and collision avoidance. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2805–2812. IEEE, 2019.

[22] Xuesu Xiao, Jan Dufek, Tim Woodbury, and Robin Murphy. Uav assisted usv visual navigation for marine mass casualty incident response. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6105–6110. IEEE, 2017.

[23] Li Qingqing, Jussi Taipalmaa, Jorge Peña Queralta, Tuan Nguyen Gia, Moncef Gabbouj, Hannu Tenhunen, Jenni Raitoharju, and Tomi Westerlund. Towards active vision with uavs in marine search and rescue: Analyzing human detection at variable altitudes. In *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 65–70. IEEE, 2020.

[24] Nicola Macoir, Jan Bauwens, Bart Jooris, Ben Van Herbruggen, Jen Rossey, Jeroen Hoebeke, and Eli De Poorter. Uwb localization with battery-powered wireless backbone for drone-based inventory management. *Sensors*, 19(3):467, 2019.

[25] Carmen Martínez Almansa, Wang Shule, Jorge Peña Queralta, and Tomi Westerlund. Autocalibration of a mobile uwb localization system for ad-hoc multi-robot deployments in gnss-denied environments. *arXiv preprint arXiv:2004.06762*, 2020.

[26] Leehter Yao, Yeong-Wei Andy Wu, Lei Yao, and Zhe Zheng Liao. An integrated imu and uwb sensor based indoor positioning system. In *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2017.

[27] Thien-Minh Nguyen, Thien Hoang Nguyen, Muqing Cao, Zhirong Qiu, and Lihua Xie. Integrated uwb-vision approach for autonomous docking of uavs in gps-denied environments. In *International Conference on Robotics and Automation (ICRA)*, pages 9603–9609. IEEE, 2019.

[28] Kun Zhang, Chong Shen, Qun Zhou, Haifeng Wang, Qian Gao, and Yushan Chen. A combined gps uwb and marg locationing algorithm for indoor and outdoor mixed scenario. *Cluster Computing*, 22(3):5965–5974, 2019.

[29] Yang Song, Mingyang Guan, Wee Peng Tay, Choi Look Law, and Changyun Wen. Uwb/lidar fusion for cooperative range-only slam. In *international conference on robotics and automation (ICRA)*, pages 6568–6574. IEEE, 2019.

[30] Thien-Minh Nguyen, Abdul Hanif Zaini, Chen Wang, Kexin Guo, and Lihua Xie. Robust target-relative localization with ultra-wideband ranging and communication. In *IEEE international conference on robotics and automation (ICRA)*, pages 2312–2319. IEEE, 2018.

[31] Jorge Peña Queralta, Li Qingqing, Fabrizio Schiano, and Tomi Westerlund. Vio-uwb-based collaborative localization and dense scene reconstruction within heterogeneous multi-robot systems. *arXiv preprint arXiv:2011.00830*, 2020.

[32] Hao Xu, Luqi Wang, Yichen Zhang, Kejie Qiu, and Shaojie Shen. Decentralized visual-inertial-uwb fusion for relative state estimation of aerial swarm. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 8776–8782. IEEE, 2020.

[33] Yang Qi, Yisheng Zhong, and Zongying Shi. Cooperative 3-d relative localization for uav swarm by fusing uwb with imu and gps. In *Journal of Physics: Conference Series*, volume 1642, page 012028. IOP Publishing, 2020.

[34] Bernhard Groβwindhager, Michael Stocker, Michael Rath, Carlo Alberto Boano, and Kay Römer. Snaploc: An ultra-fast uwb-based indoor localization system for an unlimited number of tags. In *2019 18th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 61–72. IEEE, 2019.

[35] Li Qingqing, Yu Xianjia, Jorge Peña Queralta, and Tomi Westerlund. Adaptive lidar scan frame integration: Tracking known mavs in 3d point clouds. *arXiv preprint arXiv:2103.04069*, 2021.

AN OFF-ROAD TERRAIN DATASET INCLUDING IMAGES LABELED WITH MEASURES OF TERRAIN ROUGHNESS

Gabriela Gresenz, Jules White, and Douglas C. Schmidt

{gabriela.r.gresenz, jules.white, d.schmidt}@vanderbilt.edu

Vanderbilt University
Department of Computer Science
1025 16th Avenue South, Nashville, TN 37212

ABSTRACT

This paper describes the structure and functionality of a dataset designed to enable autonomous vehicles to learn about off-road terrain using a single monocular image. This dataset includes over 12,000 images of off-road terrain and the corresponding sensor data from a global positioning system (GPS), inertial measurement units (IMUs), and a wheel rotation speed sensor. The paper also describes and empirically evaluates eight roughness labeling schemas derived from IMU z-axis acceleration for labeling the images in our dataset. These roughness labels can be used for training deep learning models to detect terrain roughness.

***Index Terms*—** Autonomous vehicles, off-road terrain, terrain roughness, deep learning, dataset

1. INTRODUCTION

Autonomous vehicle research has been a key focus in recent years, leading to the rise of increasingly autonomous vehicles operating on roadways. For example, in 2019 over 1,400 roadway autonomous vehicles across over 80 companies were in the testing stages in the U.S. alone [1]. Research on off-road terrain behavior is also important since the passenger-carrying autonomous vehicle industry eventually seeks to achieve "Level 5" autonomy, which enables entirely autonomous operation in all conditions [2]. Therefore, if a vehicle ends up in an unexpected situation—or on a route containing unmarked or unpaved terrain—the vehicle should be equipped to traverse the terrain safely.

In addition, Autonomous Ground Vehicles (AGVs) are autonomous vehicles designed to complete specific tasks without human supervision [3]. AGVs have applications in search and rescue, mining [4, 5], and planetary exploration [6]. AGVs encounter a wide range of off-road terrain that they must handle autonomously to complete their tasks safely.

Our research provides an extensive off-road terrain dataset including over 12,000 images from a monocular camera and sensor readings from a GPS, IMUs, and a wheel rotation speed sensor. We also derive eight measures of terrain roughness from the IMU z-axis acceleration readings for labeling the images in our dataset. These roughness labels can be used to train deep learning models to predict terrain roughness from monocular images.

Past research has performed semantic segmentation for understanding terrain roughness [7, 8, 9] and classification of qualitative terrain type [10, 4, 11, 12, 13]. To the best of our knowledge, however, prior work has not provided an open dataset for classifying terrain roughness as a measure of the vehicle's future kinetics from a single, monocular image.

2. RESEARCH CHALLENGES AND EXPERIMENTATION APPROACH

This section first describes the following three challenges involved in preparing an off-road terrain dataset:

- **Lack of relevant off-road terrain data**. Data collection at scale for roadway autonomous vehicles is relatively straightforward due to the vast network of roads on which humans driving vehicles equipped with sensors can travel to collect data. In contrast, there is a much smaller network of relevant off-road drivable terrain, which complicated data collection.
- **Traversing rough off-road terrain can cause an unsteady camera**, which yields images where any drivable terrain ahead is not clearly visible. Moreover, trees surround much of the drivable portions of off-road terrain. Images are therefore susceptible to poor lighting and uneven sunlight that may obstruct the image view.
- **Labeling images of upcoming drivable terrain with a single quantitative roughness metric derived from IMU z-axis acceleration readings is hard** because the length of terrain visible in an image may be unknown. As a result, determining the z-axis acceleration readings corresponding to the upcoming drivable terrain in the image is challenging. Validating that a given roughness metric effectively labels images is hard because a human may lack intuition as to how the vehicle's

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

motion will be affected by traversing this terrain, even though certain visual cues may be indicative of terrain roughness.

To address these research challenges and assist the autonomous vehicle community in making progress in off-road environments, we collected and evaluated the off-road terrain dataset described in this paper. This dataset includes eight potential roughness labeling schemas for images we collected.

The remainder of this paper is organized as follows: Section 3 describes the dataset; Sections 4 to 6 then address (1) what roughness metric should be used to label images, (2) how we selected and filtered the images in our dataset, and (3) which labeling schemas can be learned most effectively; and Section 7 presents concluding remarks and outlines future work.

3. DATASET OVERVIEW

The dataset described in this paper is available at `kaggle.com/magnumresearchgroup/offroad-terrain-dataset-for-autonomous-vehicles`. This dataset was collected in Percy Warner Park in Nashville, Tennessee, USA via a mountain bike equipped with: (1) dual GPS receivers (Garmin 830), (2) dual-high resolution IMUs (Garmin Virb Ultra), (3) a 4k 30fps camera time synchronized to both accelerometers (Garmin Virb Ultra), and (4) a wheel rotation speed sensor (Garmin Bike Speed Sensor 2). Data was collected on five different dates between late July and early October 2020.

The dataset contains sensor data and image frames extracted from videos. The videos were taken by a single monocular camera attached to the bike's handlebars. Images were extracted at one second intervals to minimize overlap between frames. The frame rate of our camera was $\sim$29.97 frames per second. It was therefore not possible to extract image frames at exactly one second intervals, so instead we found the image frame closest to each second interval.

Image frames are named by their UTC timestamps in seconds and milliseconds (e.g., "1000s100ms") and have a size of 2,160 x 3,840 pixels. We generated 12,982 images over nearly 44 miles of off-road terrain. We removed images containing sensitive information (e.g., other bikers and license plates) and images taken before or after the bike traveled the trail, resulting in 12,730 images included in our public dataset. These images are not filtered by whether they are sufficient for terrain learning (i.e., contain a clear, visible path), so researchers can access the entire range of images collected by the vehicle. Figure 1 presents some sample images from the dataset.

Sensor data is stored in a format called a "fit file," which we converted to comma-separated-value (CSV) files using tools provided by Garmin [14, 15]. We then formatted each CSV to a state-based representation, where each row contains the readings at a single timestamp and UTC timestamps so

Fig. 1. Sample Images

the data can be used alongside the image frames. The dataset contains the information described below.

1. Formatted sensor data. There is a folder for each data collection session with the following CSVs: (1) accelerometer_calibrated_split.csv, which contains the calibrated and uncalibrated acceleration readings from the accelerometer, taken $\sim$10 ms apart, (2) gyroscope_calibrated_split.csv, which contains the calibrated and uncalibrated readings from the gyroscope, taken $\sim$10 ms apart, (3) magnetometer_split.csv, which contains the uncalibrated magnetometer readings, taken $\sim$10 ms apart (4) gps.csv, which contains the vehicle's latitude, longitude, altitude, speed, heading, and velocity, taken $\sim$100 ms apart, and (5) record.csv, which contains the vehicle's latitude, longitude, distance traveled, speed, and altitude, taken 1 second apart.

2. Roughness labels for images. The following CSVs contain the eight potential roughness labels, as described in Section 4, for the subset of images valid for these labeling schemas, as described in Section 5: (1) labels_tsm1.csv contains Labels 1–4, and (2) labels_tsm2.csv contains Labels 5–8.

The accelerometer, gyroscope, magnetometer, and GPS CSV files contain system timestamps (i.e., time elapsed since the start of data collection) and calculated UTC timestamps. The GPS CSV file also contains a UTC timestamp recorded by the sensor, which may not always align with the calculated UTC timestamp due to sensor lags at certain parts of the forest.

Calibrated readings correspond directly to the x-, y-, and z-axes and are in the conventionally understood units. Our data did not contain the calibration factor for the magnetometer, so these readings remain uncalibrated. The speed and velocity readings in the GPS CSV file are GPS estimates and are significantly less accurate than the speed readings in the record CSV file, which are recorded from the wheel rotation speed sensor.

4. RESEARCH QUESTION 1: WHAT ROUGHNESS METRIC SHOULD BE USED TO LABEL IMAGES?

This section explores the derivation of our eight roughness labeling schemas based on the IMU z-axis acceleration readings for labeling images of off-road terrain.

4.1. Roughness Metric

Many studies have used z-axis acceleration to examine terrain roughness [4, 13, 8, 16]. This measure provides insight about how the vehicle's motion will be affected by traversing the upcoming terrain. Although Stavens et al. [8] standardized their measure of roughness by speed, we used a different

approach since our data did not exhibit a linear relationship between z-axis acceleration and speed. This variation likely occurred because the speeds of our vehicle (i.e., a mountain bicycle) were significantly slower than the vehicle (i.e., a car) used by Stavens et al.

Our roughness metric takes the standard deviation of a 1 second window of z-axis acceleration readings. This metric is a comprehensive measure of the terrain in the sample and is stable when the sample's mean is nonzero (such as traveling down a hill with increasing acceleration). Although our samples could reflect between 1–7 meters (since the vehicle's speed was typically between 1–7 m/s), a standard sample size was important to avoid certain samples being more susceptible to outliers than others.

We then determined which 1-second window of z-axis acceleration readings should be used to label each image. The bike traveled along particularly rough terrain, causing the angle and position of the camera to vary. The amount of upcoming terrain and its distance from the vehicle was therefore not constant across all images.

To address this issue, we explored two Terrain Sampling Methods (TSMs). *TSM 1* used a 1 second sampling of z-axis acceleration readings centered around the timestamp corresponding to 5 meters ahead of the image. *TSM 2* used a 1 second sampling of z-axis acceleration readings directly ahead of the image's timestamp.

We discretized the continuous roughness metric using each of four methods: (1) data visualization (examining the data distribution and z-axis acceleration readings alongside the continuous roughness metric), (2) k-means clustering with k = 2, (3) k-means clustering with k = 3, and (4) k-means clustering with k = 4. These methods will be referred to as original groups, k = 2 groups, k = 3 groups, and k = 4 groups, respectively. In calculating the 1 second sample for TSM 1, only 0.99 seconds of readings were included.

4.2. Labeling Images

Each image was assigned eight labels, one for each possible combination of the two methods of sampling the terrain and the four methods of discretizing the roughness metric, as shown in Table 1.

Table 1. Roughness Labeling Schemas

	Original groups	k = 2 groups	k = 3 groups	k = 4 groups
TSM 1	Label 1	Label 2	Label 3	Label 4
TSM 2	Label 5	Label 6	Label 7	Label 8

5. RESEARCH QUESTION 2: HOW DO WE SELECT AND FILTER IMAGES IN OUR DATASET?

We filtered the 12,982 images in our dataset based on sensor and visual criteria, which resulted in 7,070 images valid for Labels 1–4. To compare Labels 1–4 and Labels 5–8, we filtered the images valid for Labels 5–8 to include only images also valid for Labels 1–4, resulting in 7,061 images valid for Labels 5–8. The labeling CSVs included in our dataset do not contain two of the images used in this experiment since these images included other bikers.

We performed sensor validation to confirm the sensor readings that were either 5 meters or 3 seconds ahead of each image met the following criteria: (1) the vehicle should not be stopped, (2) sensor readings should be continuous, and (3) the calculated UTC timestamp should be within 1 second of the reported UTC timestamp. In sensor validation for Labels 1–4, we included the third criterion and GPS continuity in case other sensor readings were also affected. We did not consider these criteria for Labels 5–8 because significantly less sensor data was used to calculate these labels.

We performed visual validation to confirm that each image contained a clearly visible path. We trained an image classifier to determine which images met this criteria. Finally, we performed two rounds of manual validation to confirm the classifier's predictions.

6. RESEARCH QUESTION 3: WHICH LABELING SCHEMAS CAN BE LEARNED MOST EFFECTIVELY?

We evaluated how effectively each labeling schema could be learned by training eight different deep learning roughness classifiers, each using one of the eight labeling schemas.

6.1. Method

We split the data randomly as follows. 80% of the data was set aside for training and validation, 5% of the data was reserved for a "selection set" to select the most effectively learned labeling schemas, and 15% of the data was reserved for the testing set to provide a final evaluation of the selected models. The classes under each labeling schema were skewed, so we balanced the training-validation set by undersampling the majority classes for each labeling schema.

The image classifiers were trained in fastai [17] using transfer learning with ResNet50 [18]. We resized all images to 270 x 480 pixels to speed up training time. We used the fastai default transforms [19], which apply common image transformations to random images in the training set. We excluded the horizontal flip transform so that future work can investigate balancing classes by oversampling images in non-majority classes with a horizontal flip.

The testing and selection sets remained skewed to reflect the real-world data. We therefore evaluated the models on two metrics: overall accuracy and average accuracy by class. Average accuracy by class more heavily accounts for the model's performance on the non-majority classes, whereas overall accuracy reflects the model's performance on the actual terrain.

6.2. Analysis of Results

We trained the models and then evaluated their performance on the selection set to determine which two labeling schemas were learned most effectively, as shown in Table 2. We

Table 2. Selection Set Performance of Labeling Schemas

	TSM 1		TSM 2	
	Overall accuracy	**Avg class accuracy**	**Overall accuracy**	**Avg class accuracy**
Original groups	34.75%	36.48%	45.48%	47.72%
k = 2 groups	71.19%	71.33%	73.45%	75.06%
k = 3 groups	55.65%	46.20%	60.17%	52.30%
k = 4 groups	45.76%	35.72%	50.00%	46.27%

Table 3. Test Set Performance of Labeling Schemas

Labeling schema	Split	Overall accuracy	Avg class accuracy
Label 6	Random	69.91%	66.17%
Label 6	Chronological	70.19%	67.44%
Label 8	Random	51.32%	34.73%
Label 8	Chronological	52.92%	39.93%

first compared TSM 1 with TSM 2 by examining them side-by-side for each method of discretizing the roughness metric. TSM 2 consistently performed better than TSM 1 in both overall accuracy and average accuracy by class for each method of discretizing the roughness metric.

We then compared methods for discretizing the roughness metric, examining only TSM 2. The k = 2 groups had both the highest accuracy and highest accuracy by class, likely because the classifier had to learn only 2 categories. The jump in performance from the k = 2 groups to the k = 3 groups was significantly larger than the jump in performance from the k = 3 groups to the k = 4 groups. Likewise, the k = 4 groups provided more specific information about the upcoming terrain than the k = 3 groups. We therefore determined that the k = 4 groups were preferable compared to the k = 3 groups.

Next, we observed that the increase in overall accuracy of the k = 4 groups compared to the original groups outweighed the much smaller increase in average accuracy by class of the original groups, which indicated that the k = 4 groups were preferable. We determined that the labels learned most effectively were Labels 6 and 8, the k = 2 and k = 4 groups with TSM 2. This evaluation established a baseline for the ability to learn the various labeling schemas presented. While Labels 6 and 8 were most effectively learned using these particular architectures, other architectures may be better suited to learn other labeling schemas.

6.3. Evaluation On the Test Set

We evaluated the models corresponding to these two labels on the test set and show the results in Table 3. While we chose to extract images from the videos at 1 second intervals to minimize overlap, some images may have still contained parts of the terrain visible in chronologically consecutive images. To ensure that the roughness classifiers were learning, we further minimized potential overlap with a chronological split: the first 70% of the images in each session were used for training, the next 15% were used for validation, and the final 15% were used for testing.

We trained models using Labels 6 and 8 with the chronological split and show the results in Table 3. The chronological split models achieved comparable accuracy to the corresponding random split models and also surpassed them in both evaluation metrics. This result may have occurred because the random split allocated 5% of the data for the selection set, while these images were used in the training and validation sets for the chronological split (i.e., the chronological split had more data to train the model). This finding suggests that these models could be further improved with the addition of more training data. These results also demonstrate that the models learned to predict terrain roughness without memorizing potentially overlapping parts of terrain.

7. CONCLUDING REMARKS

This research presents a dataset for off-road terrain collected via a mountain bike. We also include eight schemas for labeling monocular images with a measure of terrain roughness derived from the IMU z-axis acceleration readings.

Based on our experiments, we identified two labeling schemas that were learned most effectively by the corresponding image classifiers: Labels 6 (TSM 2, k = 2 groups) and 8 (TSM 2, k = 4 groups). We demonstrated the performance of image classification models on these two labels, achieving 70.19% overall accuracy and 67.44% average accuracy by class for Label 6 and 52.92% overall accuracy and 39.93% average accuracy by class for Label 8.

The following are the key lessons we learned from conducting the research presented in this paper:

- **Data for off-road autonomous vehicles can be collected at scale by small, agile, and durable vehicles operated by humans.** By equipping a sturdy mountain bike with a range of sensors, we were able to gather an extensive off-road terrain dataset.
- **We can learn about the future kinetics of the vehicle as a result of upcoming terrain roughness from a single, monocular image**. While many problems in autonomous driving are being approached with expensive sensor suites, this research has shown that we can learn about terrain with a simple, low-cost sensor setup. An open research question, however, is whether these results are sufficient to control autonomous driving algorithms or if significant advancements will be necessary.

In future work we are expanding our dataset by collecting data from additional sensors, in other locations, and/or with other vehicles. We are also developing a roughness metric and a method of terrain sampling that accounts for all the visible terrain in an image. More advanced data collection equipment is needed to collect this information.

Acknowledgments

Thanks to Jiachen Xu, Shiliang Tian, and Acar Ary, who were the undergraduate researchers assisting with this project.

8. REFERENCES

[1] Darrell Etherington, "Over 1,400 self-driving vehicles are now in testing by 80+ companies across the us," Jun 2019.

[2] NHTSA, "Automated vehicles for safety," Jun 2020, Available: https://www.nhtsa.gov/technology-innovation/automated-vehicles.

[3] S. George Fernandez, K. Vijayakumar, R Palanisamy, K. Selvakumar, D. Karthikeyan, D. Selvabharathi, S. Vidyasagar, and V. Kalyanasundhram, "Unmanned and autonomous ground vehicle," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, pp. 4466, 2019.

[4] Akhil Kurup, Sam Kysar, and Jeremy P. Bos, "SVM based sensor fusion for improved terrain classification," *Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2020*, 2020.

[5] Mingliang Mei, Ji Chang, Yuling Li, Zerui Li, Xiaochuan Li, and Wenjun Lv, "Comparative study of different methods in vibration-based terrain classification for wheeled robots with shock absorbers," *Sensors*, vol. 19, no. 5, pp. 1137, 2019.

[6] NASA, "Mars 2020 Perseverance Rover," 2020, Available: https://mars.nasa.gov/mars2020.

[7] Hendrik Dahlkamp, Adrian Kaehler, David Stavens, Sebastian Thrun, and Gary Bradski, "Self-supervised monocular road detection in desert terrain," *Robotics: Science and Systems II*, 2006.

[8] David Stavens and Sebastian Thrun, "A self-supervised terrain roughness estimator for off-road autonomous driving," *arXiv:1206.6872*, 2006.

[9] Vivekanandan Suryamurthy, Vignesh Sushrutha Raghavan, Arturo Laurenzi, Nikos G. Tsagarakis, and Dimitrios Kanoulas, "Terrain segmentation and roughness estimation using rgb data: Path planning application on the centauro robot," *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)*, 2019.

[10] Yumi Iwashita, Kazuto Nakashima, Adrian Stoica, and Ryo Kurazume, "TU-Net and TDeepLab: Deep learning-based terrain classification robust to illumination changes, combining visible and thermal imagery," *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2019.

[11] Christian Weiss, Hashem Tamimi, and Andreas Zell, "A combination of vision- and vibration-based terrain classification," *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008.

[12] Chengchao Bai, Jifeng Guo, and Hongxing Zheng, "Three-dimensional vibration-based terrain classification for mobile robots," *IEEE Access*, vol. 7, pp. 63485–63492, May 2019.

[13] Christian Weiss, Holger Frohlich, and Andreas Zell, "Vibration-based terrain classification using support vector machines," *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[14] Garmin Developers, "FitCSVTool," Available: https://developer.garmin.com/fit/fitcsvtool.

[15] Garmin Developers, "FIT protocol," Available: https://developer.garmin.com/fit/protocol.

[16] Shastri Ram, *Semantic Segmentation for Terrain Roughness Estimation Using Data Autolabeled with a Custom Roughness Metric*, Ph.D. thesis, Carnegie Mellon University, 2018.

[17] fastai, "fastai v1 documentation," Available: https://fastai1.fast.ai.

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[19] fastai, "vision.transform," Available: https://fastai1.fast.ai/vision.transform.html#Data-augmentation.

A VISUAL CONTROL SCHEME FOR AUV UNDERWATER PIPELINE TRACKING

Waseem Akram and Alessandro Casavola

Department of Computer Engineering, Modeling, Electronics, and Systems (DIMES),
University of Calabria, Rende, Italy

ABSTRACT

Inspection of submarine cables and pipelines is nowadays more and more carried out by Autonomous Underwater Vehicles (AUVs) because of their low operative costs, much less than those pertaining to the traditional SHIP/ROV-based (Remotely Operated Vehicles) industrial practice, and for the improvements in their effectiveness due to technological and methodological progress in the field. In this paper, we discuss the design of a visual control scheme aimed at solving a pipeline tracking control problem. The presented scheme consists of autonomously generating a reference path of an underwater pipeline deployed on the seabed from the images taken by a camera mounted on the AUV in order to allow the vehicle to move parallel to the longitudinal axis of the pipeline so as to inspect its status. The robustness of the scheme is also shown by adding external disturbances to the closed-loop control systems. We present a comparative simulation study under Robot Operating System (ROS) to find out suitable solutions for the underwater pipeline tracking problem.

***Index Terms*—** Visual control, underwater pipeline tracking, AUV, ROS.

1. INTRODUCTION

Underwater pipelines are used as a means of transportation for oil, gas or other fluid in an underwater environment. These pipelines are prone to extreme conditions such as temperature, pressure, humidity, sea current, dust, and many more. Thus regular inspection and monitoring of these pipelines are of great importance to ensure safe transportation [1].

From past few decades, the application of autonomous underwater vehicles (AUVs) has been found in both industry and research activities as sophisticated solutions for underwater pipeline inspection and tracking. These vehicles are small in size and equipped with intelligent control, sensors, camera, and automatic navigation and localization systems. By employing camera sensors, the visual control strategies are easily implemented on these vehicles for pipeline tracking without requiring huge computations or human efforts [2].

The nature of AUV is dynamic and nonlinear, in that its behavior is coupled with highly translational and rotational dynamics. Further, these vehicles usually operate over long missions and in a dangerous and unknown environment. Strong perturbations due to sea current or actuator failure are also a source of complexity. Thus, more robust and fault-tolerant control strategies are desired. The control strategies are responsible to generate control efforts and drive the vehicle on the desired mission/trajectory. In this regard, proportional integral derivative (PID) [3], sliding mode control (SMC) [4], model predictive control (MPC) [5], linear quadratic regulator (LQR) [6] have proven to be useful control strategies for trajectory following.

The use of camera sensors and laser scanned LIDAR has allowed researchers to develop vision-based tracking systems. In particular, with the help of image processing and computer vision techniques, the desired target can be detected, located and the target trajectory obtained. This advance vision-based tracking system has substituted the traditional tracking methods such as sonar and acoustic methods and eliminated the tracking error.

In scientific literature, different studies have been carried out on underwater pipeline tracking using a vision-based control scheme. In the following, we present a brief review of some previous examples. In [7], the author proposed an image-based control scheme for AUV for solving pipeline tracking problems using fully-actuated vehicles. A plucker coordinates method is applied for an image-based feedback controller. In the experiments, both model uncertainties and external disturbances are considered. In [8], a propulsion technique is adopted for underwater pipeline tracking. For pipeline detection, morphological operations and Sobel edge detection algorithms are applied to the obtained images. Another work presented in [9] studied a vision-based approach for pipeline tracking tasks using a remotely operated vehicle (ROV) vehicle model. The work addressed image quality issues such as low-brightness and suspended materials in the underwater environment. In [10], underwater cable detection is studied using the edge classification method. In this method, edges are extracted and classified using neural networks and support vector machine algorithms. However, controller design aspects are not discussed in that work.

In this paper, a comprehensive approach is discussed to address a vision-based underwater pipeline tracking system for AUVs. The overall general architecture is shown in Figure

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

1. The proposed idea is taken from [3] and further elaborated in this work. In the latter, a vision-based tracking scheme is presented where the trajectory path is generated by using offline a sequence of images and the achieved path tested on a quadrotor. We extend the solution of that previous work by allowing the online elaboration of the images and the generated path is used to solve a pipeline tracking problem for an underwater vehicle model. The proposed solution consists of three modules. The first module is responsible to accomplish some image processing tasks. In the image processing module, we used color-based pipeline detection. The first module provides inputs to the second coordinate recovery module where a reference path is created. In the control module, a PID velocity controller is adopted. In this part, the controller generates the thruster forces based on the path tracking error. The main contribution of this work is to develop a unified approach for the vision-based underwater pipeline tracking problem using AUV in the presence of external disturbances.

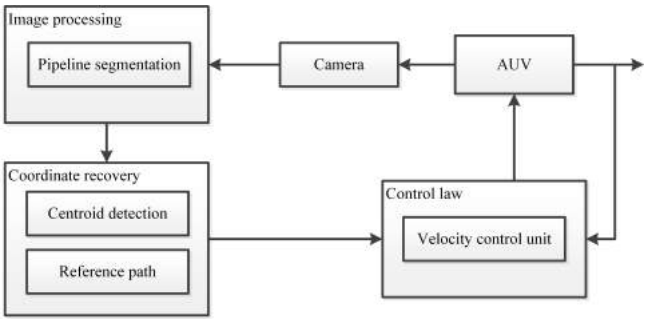


Fig. 1. Vision-based tracking scheme.

2. SIMULATION ENVIRONMENT

In this work, we used the "Unmanned Underwater Vehicle Simulator" (UUV) simulator [11]. It is a set of packages that include plugins and ROS applications that allow one to carry out simulations of underwater vehicles in Gazebo and Rviz tools.

The simulation environment consists of a seabed scene that comes along with the UUV simulator package, and an underwater vehicle used as a tracker. To simulate the underwater pipeline scenario, we have modified the seabed scene by adding a single pipeline using the blender tool. The simulation is performed on ROS melodic distribution installed on Ubuntu 18.04.4 LTS.

2.1. The rexrov model

In this work, we spawned the "rexrov-default" vehicle model available in "uuv-simulator" package of ROS [11]. The vehicle model consists of a mechanical base with a camera and additional sensing devices such as IMU (an inertial measurement unit) and LIDAR. The camera sensor is available with the vehicle model and installed at the bottom of the model that faces downward. The ROS camera easily generates image frames that have 640 pixels in width and 490 pixels in height. The "cv-bridge" plugin converts images from ROS to OpenCV and vice versa.

2.2. Initialization

The simulation initial conditions are described as follows. Let us consider an underwater pipeline placed or suspended on the surface of the ocean floor. A rexrov vehicle model is spawned on the simulation world with default positions and the same for the pipeline, as shown in Figure 2. At the start, the vehicle stays static in the sea and the camera is facing downward.

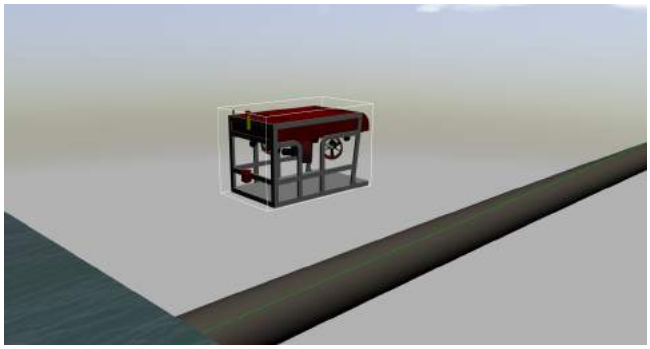


Fig. 2. Testing platform visualized in the Gazebo.

3. PROPOSED SCHEME

The vision-based underwater pipeline tracking system is achieved by using an integrated three-module-based approach. Next, each module is described as follows:

3.1. Image processing

The downward-looking camera installed on the vehicle is used to collect the image sequences to the ROS. The cv-bridge package is used to convert the obtained images from ROS to OpenCV format. Then, various image processing steps are performed to get the pipeline located in the image frame.

The following Algorithm 1 describes the basic steps taken for pipeline detection. The obtained image from the camera is shown in Figure 3. First, the image is resized to a size where the pixels and pipeline are not disturbed in the image. This is required to reduce the size of the image so the algorithm takes less time for computation. Next, the image is converted to the hue saturation value (HSV) format. The first image is used to get the HSV upper and lower values of the pipe. Then, these values are used to segment the pipe from the rest of the image. Once the pipe is successfully segmented, its pixels details are

Algorithm 1 Steps for pipe detection

Initialization:

1: **set:** lower and upper HSV of pipe in the image

Online-Phase

1: **for** $t > 0$ **do**
2: **Import** image from ROS to OpenCV
3: **Convert:** image to HSV
4: **segment:** the pipe pixels from the rest in the image
5: **save:** segmented pipe information in the list

stored in a list that will be used in the coordinate recovery module.

(a) Croped image (b) HSV image (c) masked image

Fig. 3. Image processing.

3.2. Coordinate recovery

In order to get the x and y centroid of the pipe, the "moment" method in OpenCV is used. The moment method calculates the weighted average of the image pixels of the object. In our case, here the pipe object, that was previously segmented and masked in the image, is used as input for the "moment" method. The method returns the centroid cx and cy coordinates of the pipe in the image. This is done by using the following formulas:

$$cx = \frac{1}{n}\sum_{1}^{n} x_i, cy = \frac{1}{n}\sum_{1}^{n} y_i \tag{1}$$

Here, x_i and y_i show x and y coordinates values for each pixel of the pipe and n is the total number of pixel values of the pipe. Next, the target relative x and y positions with respect to vehicle are calculated by using following formulas:

$$\begin{aligned} x &= (cx - width/2) * sensitivity, \\ y &= (cy - height/2) * sensitivity \end{aligned} \tag{2}$$

Here, the width and height of the cropped image are considered, and $sensitivity > 0$ is used to convert the pixels values to the real values as used by the simulated world frame. With the x and y values, a reference path is created during the simulation and given as input to the controller module.

3.3. Control law

In order to track the pipeline, a point-to-point reference path tracking control law is created by using the centroid information of the pipe in the image frame. The basic idea here is that the vehicle should move in order to keep the pipe in the center position of the image frame. This is achieved by setting a threshold such that the vehicle is allowed to move in the longitudinal direction along the reference path if the current error is greater than the threshold. A switching logic as shown in equation (3) is defined to accomplish this task.

$$\begin{aligned} vel.linear(x) &= vel(t), \quad |e(t)| > \ threshold \\ vel.linear(x) &= constant, \quad |e(t)| \leq \ threshold \end{aligned} \tag{3}$$

where the linear velocity of the vehicle along the x-axis $vel(t)$ is a setpoint produced by the PID control law (4) in the case the tracking error $e(t)$ is greater than the $threshold = 0.4$. The control law is responsible to reduce the tracking error $e(t)$ during the simulation. Once the tracking error is reduced below the threshold, a small positive constant number as velocity setpoint is given to the vehicle for the movement. Specifically, the thruster manager package provided by the "uuv-simulator" takes the velocity setpoint as an input that is bounded as $[-0.75, -1.3, -0.8] \leq v(t) \geq [0.75, 1.3, 0.8]$ and generates the corresponding thruster's forces. The control law is defined as:

$$vel(t) = k_p e(t) + k_i \int_0^t e(t)dt + k_d \frac{d}{dt} e(t) \tag{4}$$

Here, $k_p > 0, k_i > 0$ and $k_d > 0$ are the proportional, integral, and derivative gains respectively.

In the rest of this section we discuss the robustness of the scheme to the image blur that causes noise in the estimation of the actual position of the pipeline in the image. This is done by adding a random error in the current vehicle position throughout the simulation. The noise signal is represented by: $N \sim mN(0, 1)$ where N indicates Normal distribution with zero mean and m is a positive multiplicative term. Because of the noise, the current position is updated as follows:

$$cp(t) = cp(t) + N(t) \tag{5}$$

where cp shows the current position of the vehicle. Correspondingly, the updated error becomes

$$e(t) = tp(t) - cp(t) = tp(t) - cp(t) - N(t) \tag{6}$$

where $e(t)$ is the current error, tp target position, cp current position and N is the noise. The velocity command is updated

as follows:

$$vel(t) = k_p e(t) + k_i \int_0^t e(t)dt + k_d \frac{d}{dt} e(t) + k_p N + k_i \int_0^t N dt + k_d \frac{d}{dt} N \quad (7)$$

4. SIMULATION RESULTS

For simulation purposes, different case scenarios have been implemented as follows:

1. No-disturbance: in this case, no external disturbances are considered and handled in the controller part. This scenario is used as baseline solution.

2. PID1: in this case, $k_p = 0.5$, $k_i = 0.05$, and $k_d = 0.1$ are set along with the external disturbance 0.1

3. PID2: in this case, $k_p = 0.1$, $k_i = 0.05$, and $k_d = 0.01$ are set along with the external disturbance 0.1

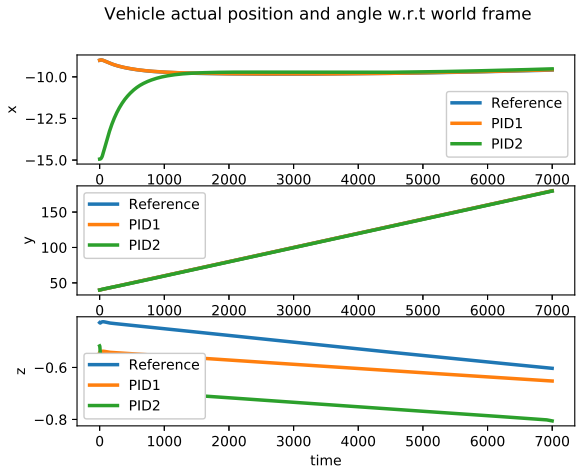


Fig. 4. Vehicle positions w.r.t simulation world frame.

Next, we discuss the simulation results. In Figure 4, the vehicle linear positions x, y and z are shown for time $t_s > 0$. The control was designed to control the x and z position for tracking purposes. In the case of PID1, the vehicle initial x position was set to -12. After a few iterations, the controller reduced the tracking error to zero. In the case of PID2, the initial x position was set to -15. It is observed that the tracking error was eliminated after a few iterations. The parameters configuration used in the PID1 case kept the angular position z close to the reference position. In Figure 5, the vehicle orientations are shown. Here, it is noticed that the vehicle stays stable through simulations. In Figure 6, the control efforts are shown that are required to control the x and z position of the vehicle. As compared with the previous work [3], the

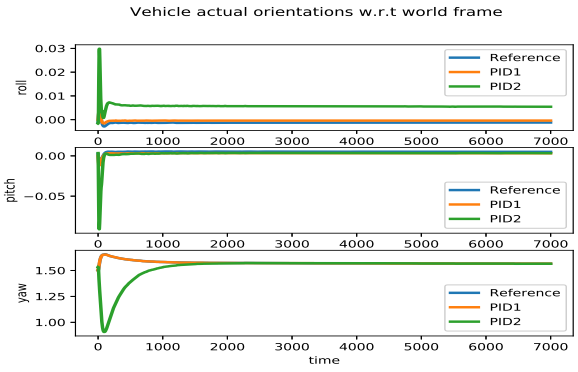


Fig. 5. Vehicle orientations w.r.t simulation world frame.

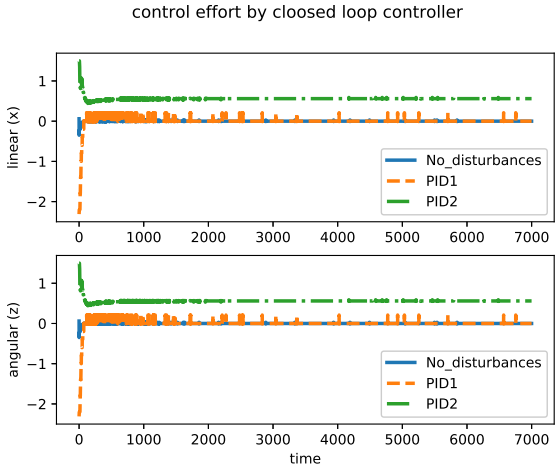


Fig. 6. Control effort of the PID controllers.

presented scheme does not require a predefined tracking path. The use of a built-in camera of AUV and in-taking few image processing steps makes it a simpler scheme for tracking problems online during run-time. Furthermore, the noise rejection capability of the scheme in the tracking position suggests implementing it in the real-world underwater environment.

5. CONCLUSIONS

The visual-based control law for underwater pipeline tracking through simulation is considered in this work. A testing platform ROS and gazebo simulator is used. A UUV simulator is adopted and modified to perform the tracking of the underwater pipeline. The proposed solution consists of three modules: image processing, coordinate recovery and PID control law. The performance of the tracking system is shown by adding and handling the external disturbance in the controller part. The system is simulated for a straight pipeline laid over the seabed. The simulation results showed a successful tracking scheme regardless of the external disturbance.

6. REFERENCES

[1] Gøril M Breivik, Sigurd A Fjerdingen, and Øystein Skotheim, "Robust pipeline localization for an autonomous underwater vehicle using stereo vision and echo sounder data," in *Intelligent Robots and Computer Vision XXVII: Algorithms and Techniques*. International Society for Optics and Photonics, 2010, vol. 7539, p. 75390B.

[2] Caoyang Yu, Xianbo Xiang, Mingjiu Zuo, and Hui Liu, "Underwater cable tracking control of under-actuated auv," in *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*. IEEE, 2016, pp. 324–329.

[3] Yusheng Wei and Zongli Lin, "Vision-based tracking by a quadrotor on ros," *Unmanned Systems*, vol. 7, no. 04, pp. 233–244, 2019.

[4] Jiehao Li, Junzheng Wang, Hui Peng, Yingbai Hu, and Hang Su, "Fuzzy-torque approximation-enhanced sliding mode control for lateral stability of mobile robot," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021.

[5] Qifan Tan, Penglei Dai, Zhihao Zhang, and Jay Katupitiya, "Mpc and pso based control methodology for path tracking of 4ws4wd vehicles," *Applied Sciences*, vol. 8, no. 6, pp. 1000, 2018.

[6] Alvaro Ortiz, Sergio Garcia-Nieto, and Raul Simarro, "Comparative study of optimal multivariable lqr and mpc controllers for unmanned combat air systems in trajectory tracking," *Electronics*, vol. 10, no. 3, pp. 331, 2021.

[7] Guillaume Allibert, Minh-Duc Hua, Szymon Krupínski, and Tarek Hamel, "Pipeline following by visual servoing for autonomous underwater vehicles," *Control Engineering Practice*, vol. 82, pp. 151–160, 2019.

[8] G Manikandan, S Sridevi, and J Dhanasekar, "Vision based autonomous underwater vehicle for pipeline tracking," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 1, pp. 2347–6710, 2015.

[9] A El-Fakdi, M Carreras, and J Batlle, "Direct policy search reinforcement learning for autonomous underwater cable tracking," *IFAC Proceedings Volumes*, vol. 41, no. 1, pp. 155–160, 2008.

[10] Mehdi Fatan, Mohammad Reza Daliri, and Alireza Mohammad Shahri, "Underwater cable detection in the images using edge classification based on texture information," *Measurement*, vol. 91, pp. 309–317, 2016.

[11] Musa Morena Marcusso Manhães, Sebastian A. Scherer, Martin Voss, Luiz Ricardo Douat, and Thomas Rauschenbach, "UUV simulator: A gazebo-based package for underwater intervention and multi-robot simulation," in *OCEANS 2016 MTS/IEEE Monterey*. sep 2016, IEEE.

SIMULTANEOUS CALIBRATION OF POSITIONS, ORIENTATIONS, AND TIME OFFSETS, AMONG MULTIPLE MICROPHONE ARRAYS

Chishio Sugiyama[1], *Katsutoshi Itoyama*[1], *Kenji Nishida*[1], *Kazuhiro Nakadai*[1,2]

[1] Dept. of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology
2-12-1 Ookayama, Meguro, Tokyo, 152-8552, JAPAN

[2] Honda Research Institute Japan Co., Ltd.
8-1 Honcho, Wako, Saitama, 351-0188, JAPAN

{sugiyama,itoyama,nishida,nakadai}@ra.sc.e.titech.ac.jp

ABSTRACT

This paper examines estimation of positions, orientations, and time offsets among multiple microphone arrays and resultant sound location. Conventional methods have limitations including requiring multiple steps for calibration, assuming synchronization between multiple microphone arrays, and necessity of a priori information, which results in convergence to a local optimal solution and large convergence time. Accordingly, we propose a novel calibration method that simultaneously optimizes positions and orientations of microphone arrays and the time offsets between them. Numerical simulations achieved accurate and fast calibration of microphone parameters without falling into a local optimum solution even when using asynchronous microphone arrays.

Index Terms— Microphone array, sound source localization, calibration, synchronization

1. INTRODUCTION

Estimating the direction of sound sources from signals observed by arrays comprising multiple microphones has been widely studied in acoustic signal processing [1]. By using multiple microphone arrays synchronized in known positions and orientations, the location of a sound source can be determined by triangulation and integration of observed sound source information. This technology has a wide range of uses, including determining the location of birds from their calls, and detecting human movement from observation of footsteps. Since it is difficult or expensive to manufacture hardware with known positions and orientations of multiple microphone arrays and to synchronize them, it is necessary to calibrate these parameters [2, 3, 4, 5].

Calibrating positions, orientations, and synchronizing small multiple microphone arrays is essentially the same as building a large microphone array. Attempts have been made to treat asynchronous distributed microphones as arrays and to calibrate their positions for synchronization [6, 7, 8]. However, since such studies investigated discrete microphones, they cannot be directly applied to microphone array calibration, because in multiple microphone arrays it is necessary to adjust not only positions and time offsets, but orientation of the microphone arrays.

Research to calibrate these parameters automatically [9, 10], was restricted by assuming synchronization or orientations were known. To estimate parameters from an unknown state, one must optimize temporal data for synchronization and spatial data for orientation. In addition, if the search space is increased to estimate from either data, it is easy to fall into a local optimum solution. This paper proposes a novel method for calibration of positions, orientations, and time offsets of multiple asynchronous microphone arrays that estimates sound source location simultaneous with calibration. We define two objective functions, time difference of arrival (TDOA) of acoustic signals among microphone arrays and direction of arrival (DOA). These objective functions avoid local optimal solutions and achieve fast convergence.

In this study, we assume that there is no synchronization between microphone arrays and investigate whether simultaneous optimization of DOA and TDOA is effective to estimate positions, orientations, and time offsets among microphone arrays. We aim to reduce the risk of convergence to a local optimal solution, simplify the algorithm, and reduce the number of sources for estimation.

2. RELATED WORK

Various proposals have been made for constructing an asynchronous distributed microphone array (ad hoc microphone array), including a method based on TDOA in a study of sound source localization [11]; one on time of arrival (TOA) [12]; one on time of flight (TOF) [13, 14]; and another on estimating from TDOA [15], but it is challenging to calibrate microphone positions.

In [16], observation time offset of each microphone was estimated. Authors in [9, 10] estimated the position and orientation of synchronized microphone arrays and location of sound sources by integrating the objective functions for DOA and TDOA and realizing the optimization of each variable. In the case of no synchronization between microphone arrays [17], a time offset was provided for each array and estimated together with position and orientation. This method included steps to first optimize placement and orientation, and then calculate coordinate scales and time offsets. Another method was to get the scale of the coordinates on assumption that the size of each microphone array was known [18]. In addition, an algorithm using RANSAC (random sample consensus) has been proposed to improve optimization based on sound source direction [19, 20].

In this study, we assume that there is no synchronization between microphone arrays and investigate whether simultaneous optimization of DOA and TDOA is effective in solving the problem of estimating position, orientation, and time offset. This method aims to reduce the risk of convergence to a local optimal solution, sim-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

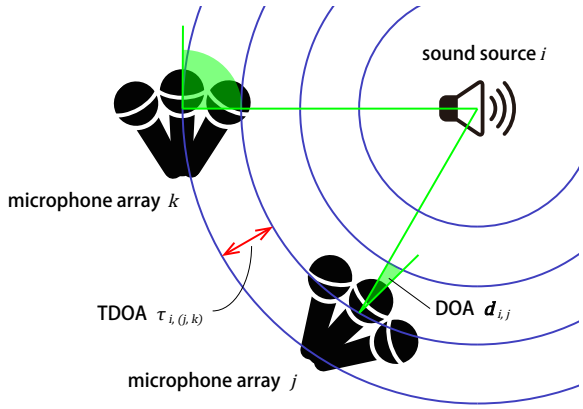


Fig. 1: A model that observes a signal from a sound source using a microphone arrays. DOA and TDOA can be calculated from each position and orientation.

plify the algorithm, and reduce the number of sources required for estimation.

3. PROPOSED METHOD

This section describes the proposed calibration method. First we review two conventional objective functions based on two different types of observations, sound source direction of arrival (DOA) and time difference of arrival (TDOA) determined from positions and orientations of microphone arrays and sound source location. Then we propose a new function by integrating these objective functions. Finally, we discuss conditions under which the proposed calibration method enables stable parameter estimation.

3.1. Signal model and conventional objective function

Let M and N be the numbers of microphone arrays and sound sources in a D-dimensional (D is 2 or 3) space, respectively. Fig. 1 shows a model of sound source direction and observation time difference. When sound source i emits an acoustic signal, a unit vector $\boldsymbol{d}_{i,j}$ representing the sound source direction as seen from the microphone array j and the observation time difference $\tau_{i,(j,k)}$ between the microphone arrays j and k are expressed by equations [17]:

$$\boldsymbol{d}_{i,j} = \mathbf{R}_j^T \frac{\boldsymbol{s}_i - \boldsymbol{a}_j}{\left\|\boldsymbol{s}_i - \boldsymbol{a}_j\right\|_2} \quad \text{and} \tag{1}$$

$$\tau_{i,(j,k)} = c\left(\left\|\boldsymbol{s}_i - \boldsymbol{a}_j\right\|_2 - \left\|\boldsymbol{s}_i - \boldsymbol{a}_k\right\|_2\right) + \delta_j - \delta_k. \tag{2}$$

where $\boldsymbol{s}_i$ and $\boldsymbol{a}_j$ are loci of sound source i and microphone array j. $\mathbf{R}_j = \mathbf{R}(\boldsymbol{\theta}_j)$ is the rotation matrix corresponding the microphone array orientation $\boldsymbol{\theta}_j$. In 2D space, $\boldsymbol{\theta}_j = \theta_{j1}$, and in 3D, $\boldsymbol{\theta}_j = (\theta_{j1}, \theta_{j2}, \theta_{j3})$, which is the ZYX Euler angles. δ_j represents the observation time offset of the microphone array j, and c is the speed of sound.

Given values of $\boldsymbol{d}_{i,j}$ and $\tau_{i,(j,k)}$, objective functions for optimizing positions, orientations, and time offsets can be written as Eqs. (3) and (4) as follows:

$$\tilde{\mathbf{S}}, \tilde{\mathbf{A}}, \hat{\boldsymbol{\theta}} = \underset{\mathbf{S},\mathbf{A},\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^{M}\sum_{j=1}^{N} D_{i,j}$$

$$D_{i,j} = \left\|\mathbf{R}_j^T(\boldsymbol{s}_i - \boldsymbol{a}_j) - \boldsymbol{d}_{i,j}\left\|\boldsymbol{s}_i - \boldsymbol{a}_j\right\|_2\right\|_2^2 \tag{3}$$

$$\hat{\mathbf{S}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\delta}} = \underset{\mathbf{S},\mathbf{A},\boldsymbol{\delta}}{\operatorname{argmin}} \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{N} T_{i,j,k}$$

$$T_{i,j,k} = \left\|\boldsymbol{s}_i - \boldsymbol{a}_j\right\|_2 - \left\|\boldsymbol{s}_i - \boldsymbol{a}_k\right\|_2 + c\,\delta_j - c\,\delta_k - c\,\tau_{i,(j,k)} \quad (j < k) \tag{4}$$

Here, $\mathbf{S} = (\boldsymbol{s}_1, ..., \boldsymbol{s}_M)$, $\mathbf{A} = (\boldsymbol{a}_1, ..., \boldsymbol{a}_N)$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_N)$, and $\boldsymbol{\delta} = (\delta_1, ..., \delta_N)$. In Eq. (3), the degree of freedom remains in the scale of the coordinates. Therefore, final coordinates $\hat{\mathbf{S}} = \alpha\tilde{\mathbf{S}}$ and $\hat{\mathbf{A}} = \alpha\tilde{\mathbf{A}}$ are obtained by multiplying the scale value $\alpha > 0$, derived from Eq. (4). The conventional method [17] (*Sequential*) solves Eq. (3) then calculates the coordinate scale α based on the observation time difference. The *Sequential* method has a high possibility of falling into a local optimum solution in the first step and uncorrected error in the second.

3.2. Proposed objective functions

To resolve this, we propose a method that alternately calculates and optimizes Eqs. (3) and (4) (*Iterate*). We combine the direction of the sound source and the objective function of the observation time difference. The optimization calculations in both Eqs. (3) and (4) find the solution by minimization making it possible to optimize for both the sound source direction and the observation time difference. The *Simultaneous* method optimizes all variables with one function by summing the minimized objective functions D and T;

$$\hat{\mathbf{S}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}} = \underset{\mathbf{S},\mathbf{A},\boldsymbol{\theta},\boldsymbol{\delta}}{\operatorname{argmin}} \sum_{i=1}^{M}\sum_{j=1}^{N}\left(D_{i,j} + \textstyle\sum_{k=1}^{N} T_{i,j,k}\right) \quad (j < k) \tag{5}$$

For the objective function of Equation (5), the combination of i, j, and k is regarded as a simultaneous equation. To reduce calculation steps, the *Integrated* method removes the common term $\left\|\boldsymbol{s}_i - \boldsymbol{a}_j\right\|_2$ from D, T;

$$\hat{\mathbf{S}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}} = \underset{\mathbf{S},\mathbf{A},\boldsymbol{\theta},\boldsymbol{\delta}}{\operatorname{argmin}} \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{N} I_{i,j,k}$$

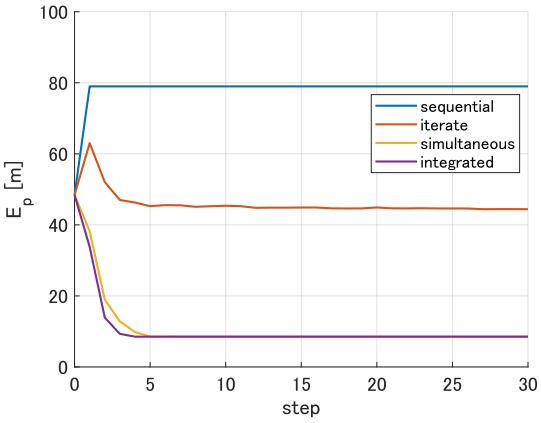
$$I_{i,j,k} = \left\|\mathbf{R}_j^T(\boldsymbol{s}_i - \boldsymbol{a}_j) - \boldsymbol{d}_{i,j}\left(T_{i,j,k} - \left\|\boldsymbol{s}_i - \boldsymbol{a}_j\right\|_2\right)\right\|_2^2 \quad (j \le k) \tag{6}$$

We verify all four methods of estimation in 2D space, and *Simultaneous* and *Integrated* in 3D. To determine the solution, we fix coordinate $\boldsymbol{s}_1$ to origin $\boldsymbol{O}$ and angle of rotation $\boldsymbol{\theta}_1$ and the observation time offset δ_1 to 0[deg] and 0[s], respectively. Optimization calculations use the interior point method to find the solution.

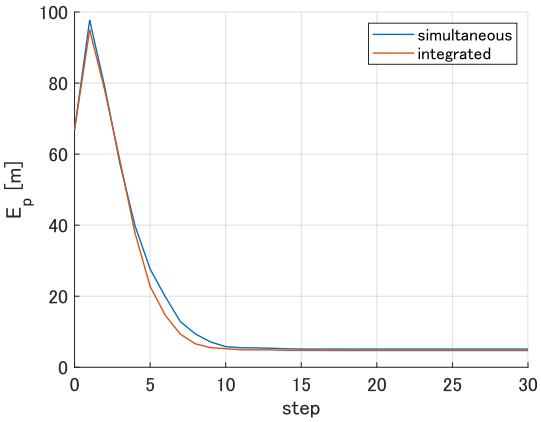
3.3. Number of sound sources required for estimation

When estimating positions of microphone arrays and locations of sound sources by the above methods, acoustic signals are observed from sound sources existing at different coordinates. The number of sound sources required for estimation is determined by the number of variables to be estimated and the number of optimization formulas.

The numbers of variables are $2M$ for source coordinates, $2(N-1)$ for microphone array coordinates, $N-1$ for orientation, and $N-1$ for time offset in 2D space. In 3D, $3M$, $3(N-1)$, $3(N-1)$, and $N-1$ respectively. Count the number of linearly independent combinations of i, j, and k in an error-free optimization calculation.



(a) Average error in estimation of positions of microphone arrays and locations of sound sources (2D). The *Simultaneous* and *Integrated* methods that optimize DOA and TDOA with a single objective function are highly accurate.



(b) Average error in estimation of positions of microphone arrays and locations of sound sources (3D). The method estimated correctly in 2D is also highly accurate in 3D.

Fig. 2: Localization error

Then, (3) has MN, (4) has $M(N-1)$, (5) and (6) have $MN + M(N-1) = M(2N-1)$ combinations.

For a good estimate, make the number of equations greater than or equal to the number of variables. For example, when $N = 3$ in 2D space, number of sound sources M is selected so that $M \geq 6$ holds for *Sequential* and *Iterate*, and $M \geq \frac{8}{3}$ holds for *Simultaneous* and *Integrated*.

4. EVALUATION

4.1. Method performance comparison

A numerical simulation was conducted to verify the effectiveness of the proposed method. We generated 100 patterns of values under the condition of 7 sound sources and 3 microphone arrays placed in $120 \times 120[\mathrm{m}^2]$ (2D) or $120 \times 120 \times 120[\mathrm{m}^3]$ (3D) space with uniformly distributed random values. The microphone array orientation values were uniformly distributed, and time offset values with a normal distribution $\mathcal{N}(0, 10^{-4})[\mathrm{s}]$ were generated.

Table 1: Shows the time [s] taken to calculate optimization. *Integrated* is the shortest time in both 2D and 3D.

Method	2D	3D
Sequential	1.71	–
Iterate	4.36	–
Simultaneous	1.97	9.42
Integrated	1.54	8.32

Sound source direction $\boldsymbol{d}$ and observation time difference τ were calculated from the generated values and used as correct data. $\mathcal{N}(0, 4)[\mathrm{deg}]$ error was added to the localization of the microphone array and observation resolutions were set to $1[\mathrm{deg}]$. The error of observation time differences was set to $\mathcal{N}(0, 10^{-6})[\mathrm{s}]$, and sampling frequencies to $16[\mathrm{kHz}]$.

$$E_p = \frac{1}{M+N}\left(\sum_{i=1}^{M}\left\|\hat{\boldsymbol{s}}_i - \check{\boldsymbol{s}}_i\right\|_2 + \sum_{j=1}^{N}\left\|\hat{\boldsymbol{a}}_j - \check{\boldsymbol{a}}_j\right\|_2\right) \tag{7}$$

Fig. 2 shows transition of localization error E_p for each optimization method. $\hat{\cdot}$ is the estimated value and $\check{\cdot}$ is the true value. The average error for 100 different placements is recorded as one step for each fixed evaluation count of the optimization calculation.

The *Sequential* method converges to a local optimal solution in the first step, although the error of *Iterate* estimation is smaller than that of *Sequential*, it is classified as a local optimal solution. On the other hand, both the *Simultaneous* and *Integrated* methods had an error of about $8[\mathrm{m}]$. This was considered to be the result of accurate estimation, considering that an error was added to the sound source direction and observation time difference.

Under simulation conditions, it is assumed sound source direction and generation time can be measured with high accuracy. Since accuracy deteriorates with reverberations in the real world, accuracy of estimation by actual observation is a future task.

4.2. Orientations and time offsets estimation accuracy

We investigated accuracy of microphone array orientation and time offset estimation in 3D. The acquisition resolution of the sound source direction of the array was changed 0-5[deg], and a simulation was performed every $1[\mathrm{deg}]$ with 100 patterns arranged. Other settings were kept as in the previous experiment. Since the reference microphone array error was 0, accuracy was calculated from values of the other two arrays. The results are shown in Figs. 3 and 4.

Orientation error E_d is defined by

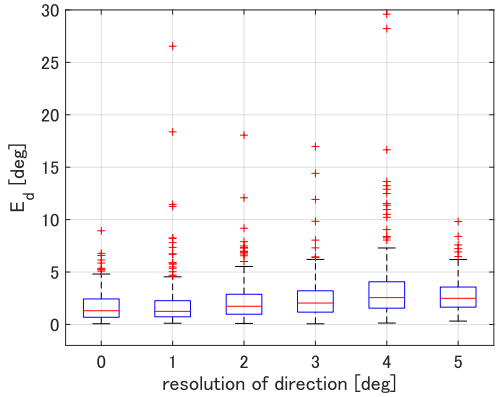
$$E_d = \arccos\left(\frac{\hat{\mathbf{R}}_j\boldsymbol{v} \cdot \check{\mathbf{R}}_j\boldsymbol{v}}{\|\boldsymbol{v}\|_2^2}\right) \ (j = 2, \ldots, N)\,. \tag{8}$$

The vector $\boldsymbol{v} = [1, 1, 1]^T$ is rotated by rotation matrix $\mathbf{R}$, and accuracy of the direction is calculated from its cosine similarity. The time offset is evaluated by RMSE (root mean square error) for each placement pattern.

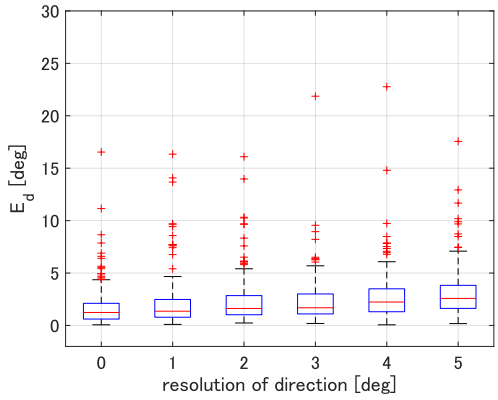
Accuracy did not differ between methods. Directional error was $< 10[\mathrm{deg}]$ and time offset error $< 10[\mathrm{ms}]$. Both error values increased as resolution of sound source orientation decreased.

4.3. Processing time

Table 1 shows average time to complete estimation for one placement in the simulation in 4.1.

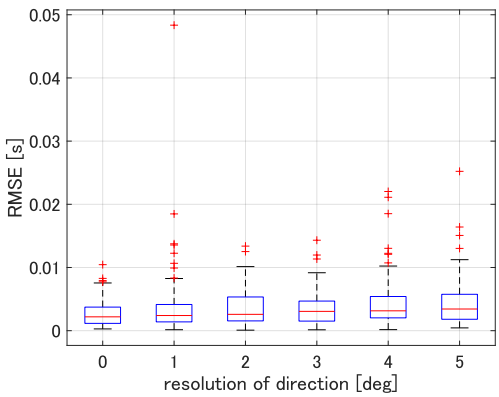


(a) simultaneous

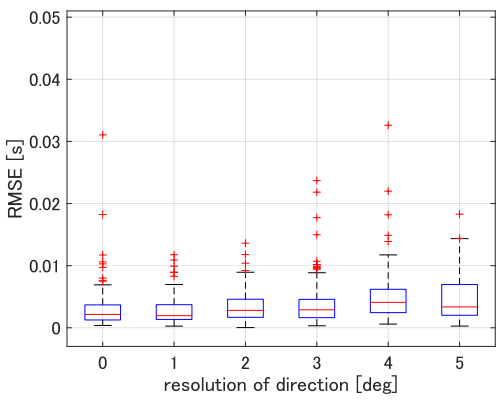


(b) integrated

Fig. 3: Evaluation of orientation error of microphone arrays with E_d (3D). As resolution of sound source direction decreased, error slightly increased.



(a) simultaneous



(b) integrated

Fig. 4: RMSE of time offset error for each arrangement (3D). There was no difference in accuracy between the two methods.

In 2D, *Iterate* takes the longest time while *Integrated* takes less than *Sequential* which quickly converges to local optimal solution. Comparing *Simultaneous* and *Integrated* with high estimation accuracy, *Integrated* was calculated in a short time in both 2D and 3D. However, considering that 3D estimation takes 8–9[s], it is necessary to further reduce calculation time to introduce it as an online estimation system.

5. CONCLUSION

We described an optimization method to estimate sound source locations, and positions, orientations, and time offsets of microphone arrays, using their observations of acoustic signals from multiple sources. Calculation of the optimal solution using a simple algorithm showed the method of simultaneously referencing sound source direction with respect to the microphone array and acoustic signal arrival time differences among microphone arrays yields more accurate estimation than optimizing them independently. Estimation accuracy of *Simultaneous* and *Integrated*, was almost the same. However, since the calculation time of *Integrated* was shorter, the method proposed in this paper is *Integrated*.

Performance was evaluated by simulation on the assumption of little error in input data. A future task is performance evaluation in a real setting including reverberation. For use on robots and drones that observe in motion, a method to estimate in a shorter time in 3D space in necessary.

Acknowledgements This work was supported by JSPS KAKENHI Grant No. JP19K12017, JP19KK0260, and JP20H00475.

6. REFERENCES

[1] F. Asano, *Array signal processing for acoustics –Localization, tracking and separation of sound sources–*, Corona Publishing Co., Ltd., 2011, (in Japanese).

[2] T. Yamada, K. Itoyama, K. Nishida, , and K. Nakadai, "3D sound source tracking for drones using direction likelihood integration," in *JSAI Technical Report, SIG-Challenge-057-02*, 2020, (in Japanese).

[3] D. Gabriel, R. Kojima, K. Hoshiba, K. Itoyama, K. Nishida, and K. Nakadai, "2D sound source position estimation using microphone arrays and its application to a VR-based bird song analysis system," *Advanced Robotics*, vol. 33, no. 7-8, pp. 403–414, 2019.

[4] S. D. Valente, M. Tagliasacchi, F. Antonacci, P. Bestagini, A. Sarti, and S. Tubaro, "Geometric calibration of distributed microphone arrays from acoustic source correspondences," in *2010 IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2010, pp. 13–18.

[5] A. Plinge and G. A. Fink, "Geometry calibration of distributed microphone arrays exploiting audio-visual correspondences," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, 2014, pp. 116–120.

[6] S. Thrun, "Affine structure from sound," in *Advances in Neural Information Processing Systems*, 2006, pp. 1353–1360.

[7] K. Dan, K. Itoyama, K. Nishida, and K. Nakadai, "Calibration of a microphone array based on a probabilistic model of microphone positions," in *Trends in Artificial Intelligence Theory and Applications. Artificial Intelligence Practices (IEA/AIE)*, 2020, pp. 614–625.

[8] H. Miura, T. Yoshida, K. Nakamura, and K. Nakadai, "SLAM-based online calibration of asynchronous microphone array for robot audition," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 524–529.

[9] A. Plinge and G. A. Fink, "Geometry calibration of multiple microphone arrays in highly reverberant environments," in *2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2014, pp. 243–247.

[10] A. Plinge, G. A. Fink, and S. Gannot, "Passive online geometry calibration of acoustic sensor networks," *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 324–328, 2017.

[11] A. Canclini, F. Antonacci, A. Sarti, and S. Tubaro, "Acoustic source localization with distributed asynchronous microphone networks," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 2, pp. 439–443, 2013.

[12] S. T. Birchfield, "Geometric microphone array calibration by multidimensional scaling," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2003, vol. 5, pp. 157–160.

[13] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 660–673, 2012.

[14] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 2597–2600.

[15] Y. Kuang and K. Åström, "Stratified sensor network self-calibration from tdoa measurements," in *21st European Signal Processing Conference (EUSIPCO 2013)*, Sep. 2013, pp. 1–5.

[16] P. Pertilä, M. S. Hämäläinen, and M. Mieskolainen, "Passive temporal offset estimation of multichannel recordings of an ad-hoc microphone array," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 11, pp. 2393–2402, 2013.

[17] S. Woźniak and K. Kowalczyk, "Passive joint localization and synchronization of distributed microphone arrays," *IEEE Signal Processing Letters*, vol. 26, no. 2, pp. 292–296, 2019.

[18] F. Jacob, J. Schmalenstroeer, and R. Haeb-Umbach, "DOA-based microphone array postion self-calibration using circular statistics," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2013, pp. 116–120.

[19] J. Schmalenstroeer, F. Jacob, R. Haeb-Umbach, M. H. Hennecke, and G. A. Fink, "Unsupervised geometry calibration of acoustic sensor networks using source correspondences," in *12th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2011.

[20] F. Jacob, J. Schmalenstroeer, and R. Haeb-Umbach, "Microphone array position self-calibration from reverberant speech input," in *2012 12th International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2012, pp. 1–4.

IMPROVED AND EFFICIENT INTER-VEHICLE DISTANCE ESTIMATION USING ROAD GRADIENTS OF BOTH EGO AND TARGET VEHICLES

Muhyun Back[1], *Jinkyu Lee*[1], *Kyuho Bae*[2], *Sung Soo Hwang*[1,*], *Il Yong Chun*[3,*]

[1]School of Computer Science and Electrical Engineering, Handong Global University
[2]Stradvision Inc.
[3]Department of Electrical and Computer Engineering, University of Hawai'i at Mānoa

ABSTRACT

In advanced driver assistant systems and autonomous driving, it is crucial to estimate distances between an ego vehicle and target vehicles. Existing inter-vehicle distance estimation methods assume that the ego and target vehicles drive on a same ground plane. In practical driving environments, however, they may drive on different ground planes. This paper proposes an inter-vehicle distance estimation framework that can consider slope changes of a road forward, by estimating road gradients of *both* ego vehicle and target vehicles and using a 2D object detection deep net. Numerical experiments demonstrate that the proposed method significantly improves the distance estimation accuracy and time complexity, compared to deep learning-based depth estimation methods.

Index Terms— Inter-vehicle distance estimation, Autonomous driving, ADAS, Visual odometry

1. INTRODUCTION

There have been many advances in advanced driver assistance systems (ADAS) and autonomous driving technologies. To achieve safe driving, it is important to understand driving environments, such as the existence of obstacles around the vehicle. Estimating the distances between an ego vehicle to other objects on the road is essential to maintain a safe distance to other vehicles and avoid obstacles, etc.

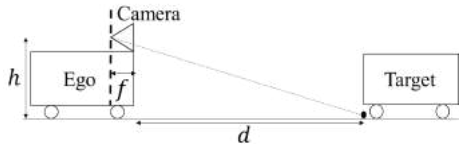
Many existing inter-vehicle distance estimation methods use LiDAR or RADAR sensors [1, 2] with the high costs, where this paper refers distance estimation between an ego vehicle and target vehicle(s) as inter-vehicle distance estimation. An alternative to such methods using high-cost sensors is using a monocular camera in distance estimation. The conventional monocular camera-based distance estimation methods use 2D image processing and physical information of a target vehicle [3, 4, 5]. Specifically, the methods need to know the rear width of a target vehicle or the height of the ego vehicle and such requirements limit the practical use of the distance estimation techniques. Alternatively, deep learning-based depth estimation methods have been proposed [6, 7]. The methods estimate a depth of entire scenes, but does not specifically estimate a distance between vehicles. These methods have generalization risks, and such risks become problematic if driving environments captured in camera(s) are different between training and test. In addition, the deep learning-based depth estimation methods are computationally expensive, and it may be challenging to use them combined with object detection deep neural networks (DNNs) in real-time applications, e.g., autonomous driving. Recently, [8] proposed a camera projection-based inter-vehicle estimation method aided with a 3D object detection DNN.

Camera model-based distance estimation methods assume that ego and target vehicles consistently drive on a same ground plane [8, 9, 10]. In practice, however, they may not be on a same ground plane: either ego or target vehicle may drive uphill or downhill, and slope of a road forward may change. An alternative is to consider road gradients, where road gradient refers to the steepness of a road. A flat road is said to have a road gradient of 0°; an uphill is said to have a positive road gradient; a downhill is said to have a negative road gradient. The recent distance estimation method [11] considers road gradient of an ego vehicle using an inertial measurement unit (IMU). Nonetheless, this method does not consider road gradient of a target vehicle, and it might become inaccurate, if road gradient of a target vehicle changes due to slope changes of a road forward.

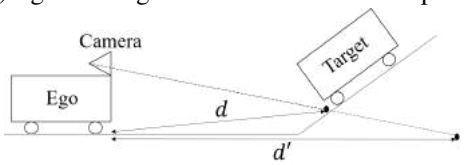
This paper proposes a distance estimation framework that uses estimated road gradients of both ego vehicle and target vehicles. To estimate road gradient of an ego vehicle, we use the general property that vehicles on a road are parallel to a ground plane and vehicle pose and road gradient changes are closely related. We estimate ego vehicle's road gradient based on changes of vehicle poses encapsulated in camera rotation matrix at different time points, where camera rotation matrices are estimated by monocular visual odometry [12]. This is different from [11] that uses an IMU in estimating camera rotation matrix. To estimate road gradient of a target vehicle, we

*Corresponding Authors

This research is supported in part by StradVision, seed grants from the University of Hawai'i at Mānoa, and Ingeborg v.F. McKee Fund of the Hawai'i Community Foundation.



(a) Ego and target vehicles are on a same plane.



(b) Ego and target vehicles are on different planes.

Fig. 1. Schematic diagrams of distance estimation using a pinhole camera.

first determine whether ego and target vehicles are driving on a same ground plane. Specifically, we compare a position of target vehicle estimated by deep-learning based object detection and a vanishing line calculated from the camera rotation matrix estimated above. Using the comparison results, we estimate a road gradient of a target vehicle. If ego and target vehicles have sufficiently different road gradients, we adjust road gradient of an ego vehicle by some proper amount. Numerical experiments with six datasets show that the proposed method significantly improves the distance estimation accuracy and time complexity, compared to deep learning-based depth estimation methods, DORN [6] and Monodepth2 [7].

2. BACKGROUND

This section briefly reviews the pinhole camera model-based distance estimation method. The method assumes that ego and target vehicles drive on a same ground plane; Fig. 1(a). The distance between ego and target vehicles, d, can be estimated by the following formula [11, (3)]:

$$d = \frac{f \cdot h}{\delta_y \cdot (u - v)} \tag{1}$$

where f is the camera focal length, h is the camera height, δ_y is the physical size of a pixel along the y-axes in the image domain, u is a position of the bottom of a target vehicle in an image, and v is a position of the vanishing line in an image.

The parameters f, h, and δ_y are obtained in camera calibration and do not change in driving. The variable v is affected by a 3×3 camera rotation matrix $\mathbf{R}$ [13, Ch. 8], and inaccurate v can cause errors in calculating d. This camera rotation matrix $\mathbf{R}$ changes according to road changes such as slope and direction changes. [11] uses an IMU to keep updating $\mathbf{R}$. The variable u also changes while driving, and we observed that object detection DNNs accurately estimate it.

Fig. 1(b) shows a schematic diagram of a pinhole camera when ego and target vehicles drive on different ground planes. If a target vehicle's road gradient is not considered, method

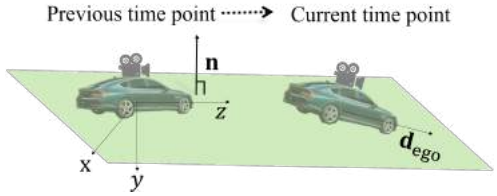


Fig. 2. Changes of an ego vehicle's pose between previous and current time points.

(1) incorrectly estimates the distance between ego and target vehicles as d' instead of d.

3. PROPOSED METHOD

The proposed method uses video frames and deep learning-based object detection results in estimating road gradients of ego and target vehicles. We estimate road gradient of an ego vehicle using monocular visual odometry with video frames. To estimate road gradient of target vehicles, we use a 2D bounding box around a target vehicle that is estimated from deep learning-based object detection. To measure a ground plane difference between ego and target vehicles, we convert a road gradient of an ego vehicle to a vanishing line, and compare the position of calculated vanishing line, and that of the center position of a bounding box around target vehicles. If this measure is less than or equal to some threshold, we assume that ego and target vehicles drive on different ground planes; otherwise, they drive on a same ground plane. If we decide that they drive on different ground planes, we adjust road gradient of an ego vehicle and use adjusted one to estimate distance between ego and target vehicles. If we decide that they drive on a same ground plane, road gradient of an ego vehicle is used to estimate inter-vehicle distances. For simplicity, the section describes the proposed distance estimation method between an ego vehicle and a target vehicle.

3.1. Road gradient estimation of an ego vehicle

We estimate road gradient of an ego vehicle using monocular visual odometry. The monocular visual odometry is defined on the x-, y- and z-axes [12] – see Fig. 2 – and calculates a camera rotation matrix difference $\mathbf{\Delta R} \in \mathbb{R}^{3\times3}$ between each frame and a starting point. We assume that at a current time point, an ego vehicle drives in the direction of an unit vector $\mathbf{z}$. The proposed method calculates road gradient of an ego vehicle, θ, by calculating an angle between the normal vector $\mathbf{n}$ to the ground plane, i.e., xz-plane in Fig. 2, that is updated at a previous time point, and current direction $\mathbf{z}$.

Using calculated $\mathbf{\Delta R}$ via the monocular visual odometry, we calculate the normal vector to the xz-plane at a previous time point as follows:

$$\mathbf{n} = \mathbf{\Delta R}_{\text{prev}} \cdot \mathbf{n}_0, \quad \mathbf{n}_0 = \begin{bmatrix} 0 \\ -\cos(\theta_0) \\ -\sin(\theta_0) \end{bmatrix}, \tag{2}$$

where $\mathbf{\Delta R}_{\text{prev}}$ is the camera rotation matrix difference between previous and initial time points, θ_0 is a camera angle in the yz-plane. We calculate the current direction of an ego by

$$\mathbf{d}_{\text{ego}} = \mathbf{\Delta R}_{\text{curr}} \cdot \mathbf{d}_{\text{ego},0}, \quad \mathbf{d}_{\text{ego},0} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{3}$$

where $\mathbf{\Delta R}_{\text{curr}}$ is the camera rotation matrix difference between current and initial time points, $\mathbf{d}_{\text{ego},0}$ is the default direction of an ego vehicle. Using (2) and (3), we calculate road gradient of an ego vehicle by finding an angle between $\mathbf{n}$ and $\mathbf{d}_{\text{ego}}$:

$$\theta = \frac{\pi}{2} - \arccos\left(\mathbf{n}^\top \mathbf{d}_{\text{ego}}\right). \tag{4}$$

noting that both $\mathbf{n}$ and $\mathbf{d}_{\text{ego}}$ are unit vectors.

To properly calculate the road gradient of an ego vehicle at a current time point, there must be some time interval between previous and current time points. If there exist no time interval, then $\mathbf{\Delta R}_{\text{prev}} = \mathbf{\Delta R}_{\text{curr}}$ and thus, $\theta = 0$ does not change over time, assuming that $\theta_0 = 0$ (i.e., no camera rotation in the yz-plane). We assume that an ego vehicle drives at 60 km/h, and set the time interval between previous and current time points as 1 sec.

3.2. Road gradient estimation of a target vehicle

To estimate road gradient of a target vehicle, we use road gradient of an ego vehicle, θ in (4), and a 2D bounding box around a target vehicle that is obtained via deep learning-based object detection. The variable v in (1) can be calculated with θ in (4) using the following formula [14]:

$$v = c_y - \tan\theta \cdot f, \tag{5}$$

where c_y is a y-coordinate of the principal point or image center in the image domain (we set the most left corner pixel location as $(0, 0)$), and f is given in (1). The parameter c_y is obtained in camera calibration and does not change in driving.

We observed that when ego and target vehicles drive on different ground planes, the (bounding box) center of a target vehicle is formed in a deviated location from its usual position formed by those driving on a same ground plane. When ego and target vehicles drive on a same ground plane, the center of a target vehicle is positioned slightly below a vanishing line; see Fig. 3(a). When a target vehicle drives on a low uphill slope, the center of a target vehicle gets closer to a vanishing line. When a target vehicle drives on a steep uphill slope, the center of a target vehicle can be positioned above the vanishing line; see Fig. 3(b). In case of a target vehicle driving on a downhill slope, the center of a target vehicle moves downward from the vanishing line.

We estimate road gradient changes of a target vehicle and adjust road gradient of an ego vehicle by using the aforementioned positional relationship between the vanishing line and the center of a target vehicle. To estimate road gradient changes of a target vehicle, we calculate the pixel coordinate difference between the vanishing line position, v in (5), and the center position of a target vehicle, b_y, in a y-coordinate:

(a) Ego and target vehicles drive on a same ground plane.

(b) Ego and target vehicles drive on different ground planes.

Fig. 3. Positions of estimated vanishing line and center of a bounding box around a target vehicle. Green solid lines denote the vanishing line of a driving scene, and yellow dashed lines denote the center position of 2D bounding box around a target vehicle.

$$\Delta_y = b_y - v. \tag{6}$$

When ego and target vehicles drive on a same ground plane, the value of Δ_y is positive. When a target vehicle starts driving on uphill slope, Δ_y becomes smaller, and if the slope becomes steep, Δ_y eventually becomes negative. We mainly investigate the case that a target vehicle drives on uphill slope (the majority of frames in this paper include either this case or the case that ego and targets vehicles drive on a same ground plane; see Section 4.1). We adjust the road gradient of an ego vehicle θ by considering ground plane difference between ego and target vehicles – in Fig. 1(b), to obtain d instead of d':

$$\theta \leftarrow \begin{cases} \theta + \alpha_1, & \Delta_y \leq 0, \\ \theta + \alpha_2, & \Delta_y \leq -10, \\ \theta + \alpha_3, & \Delta_y < -20, \\ \theta, & \text{otherwise}, \end{cases} \tag{7}$$

where $0 < \alpha_1 < \alpha_2 < \alpha_3$ are tunable angle adjustment parameters. The otherwise condition in (7) implies that we do not adjust θ if the road gradient difference between ego and target vehicles is small, i.e., the difference between d and d' in Fig. 1(b) is negligible.

4. RESULTS AND DISCUSSIONS

4.1. Experimental setup

We compare the proposed method with DORN [6] and Monodepth2 [7], depth estimation DNN trained in a supervised and unsupervised way, respectively. For the comparisons, we use the video the sequence of KITTI validation split with 1024×368 frame size [15] (the validation split has sufficiently long depth estimation video) and five video sequences provided by Stradvision (SV) with 1920×1080 frame size [16]; we refer them as SV Sequences #1, . . . ,#5. The KITTI video sequence includes cases that ego and target vehicles drive on uphill, downhill and curved road, and they drive on *almost* same ground plane in all cases. Each SV video sequence has different driving environments: in SV Sequences #1 and #2, ego and target vehicles drive on a same flat road and a same downhill slope, respectively; in SV Sequence #3, ego and target vehicles drive and curve on a same flat road; SV Sequence #4 includes speed bumps and curved flat roads; SV Sequence #5 includes the case that a target vehicle drive on steep uphill slope. The KITTI and SV video sequences include many target vehicles and a single target vehicle, respectively. In the KITTI video sequence, the proposed method estimates distance between an ego vehicle and *all* target vehicles. The number of frames of KITTI and SV Sequences is 840 and 500, respectively. The KITTI video sequence obtains the ground truth inter-vehicle distances with LiDAR data. The SV video sequences obtain those with differential GPS.

In proposed (7), we set the angle adjustment parameters as follows: $\alpha_1 = 3^\circ$, $\alpha_2 = 5^\circ$ and $\alpha_3 = 6^\circ$. We used the SV object detection software [16] to estimate the 2D bounding box of each target vehicle. We used pre-trained DORN and Monodepth2 networks that were trained by KITTI dataset [6, 7]. These methods consider estimated depth at the bounding box center of target vehicle(s) (obtained by the SV object detection software) as inter-vehicle distance. The IMU data does not exist in the KITTI and SV datasets, so comparisons with [11] are omitted.

We evaluated the inter-vehicle distance estimation accuracy by the most conventional error metric in the distance estimation and depth estimation applications, root-mean-squared-error (RMSE in m). In measuring computing times (in secs.) of the proposed method and DORN & Monodepth2, we used Intel Core I7-8700 CPU with 3.20GHz and 64GB RAM, and NVIDIA GeForce GTX 1060 with 6GB, respectively. The measured execution time of the proposed method additionally includes the visual odometry computing time. All the methods used the 2D object detection software provided by SV, so we did not measure its computation time.

4.2. Result and discussion

In all experiments, the proposed method significantly improves the inter-vehicle distance estimation accuracy compared to the deep-learning based depth estimation methods, DORN [6] and Monodepth2 [7]. See Table 1. The SV Sequence #1–#3 results in Table 1 show that when ego and target vehicles drive on a same ground plane, the proposed method significantly improves the inter-vehicle distance estimation of an ego vehicle compared to DORN and Monodepth2, regardless of driving environments. The SV Sequence #4–#5 results in Table 1 show that when ego and target vehicles drive on different ground planes, the proposed method significantly improves the inter-vehicle distance estimation compared to DORN and Monodepth2. Before updating road gradients of an ego vehicle (see Section 3.2), RMSE values are 9.6 (m) and 13.86 (m) for SV Sequences #4 and #5, respectively. Comparing the results with those in Table 1 implies that the proposed method in Section 3.2 is crucial to accurately estimate inter-vehicle distances when ego and target vehicles drive on different ground planes. Experiments with the KITTI dataset has higher errors than other datasets, because all methods estimate distance between an ego vehicle and many target vehicles.

Table 1. RMSE values (m) of different inter-vehicle distance estimation methods

	Ours	DORN	Monodepth2
KITTI	**7.63**	8.35	11.5
SV Seq. #1	**1.72**	8.88	5.93
SV Seq. #2	**3.49**	7.22	9.77
SV Seq. #3	**1.25**	2.86	4.63
SV Seq. #4	**4.3**	6.57	8.7
SV Seq. #5	**5.95**	9.03	10.35

Table 2. Averaged processing time (secs.) per frame of different methods

	Ours	DORN	Monodepth2
KITTI	**0.03**	3	0.06
SV Seq. #1–5	**0.04**	12	0.36

Table 2 shows that in all experiments, the proposed method is consistently faster than depth estimation DNNs, DORN and Monodepth2. We expect that the proposed method becomes more efficient than the other depth estimation methods as the resolution of video increases.

5. CONCLUSION

The proposed method achieves accurate and fast inter-vehicle distance estimation by estimating road gradient of *both* ego and target vehicles. The proposed method is expected to be useful for autonomous driving and ADAS in practical driving environments, where road gradients of ego and target vehicles change over time.

We will investigate an advanced DNN that updates ego vehicle's road gradient θ based on v in (5) and b_y in (6).

6. REFERENCES

[1] R. Thakur, "Scanning LIDAR in Advanced Driver Assistance Systems and Beyond: Building a road map for next-generation LIDAR technology," *IEEE Consumer Electronics Magazine*, vol. 5, no. 3, pp. 48–54, August 2016.

[2] G. Hakobyan and B. Yang, "High-performance automotive radar: A review of signal processing algorithms and modulation schemes," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 32–44, September 2019.

[3] E. Dagan, O. Mano, G. P. Stein, and A. Shashua, "Forward collision warning with a single camera," in *Proc. IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004, pp. 37–42.

[4] G. Kim and J. S. Cho, "Vision-based vehicle detection and inter-vehicle distance estimation for driver alarm system," *Optical Review*, vol. 19, no. 6, pp. 388–393, December 2012.

[5] A. Ali, A. Hassan, A. R. Ali, H. U. Khan, W. Kazmi, and A. Zaheer, "Real-time vehicle distance estimation using single view geometry," in *Proc. IEEE Winter Conference on Applications of Computer Vision*, Snowmass Village, CO, March 2020, pp. 1111–1120.

[6] H. Fu, M. Gong, C. Wang, K. Batmanghelich, and D. Tao, "Deep ordinal regression network for monocular depth estimation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, June 2018, pp. 2002–2011.

[7] C. Godard, O. M. Aodha, M. Firman, and G. Brostow, "Digging into self-supervised monocular depth prediction," in *Proc. IEEE International Conference on Computer Vision*, Seoul, South Korea, October 2019, pp. 3827–3837.

[8] T. Zhe, L. Huang, Q. Wu, J. Zhang, C. Pei, and L. Li, "Inter-vehicle distance estimation method based on monocular vision using 3D detection," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4907–4919, March 2020.

[9] G. P. Stein, O. Mano, and A. Shashua, "Vision-based ACC with a single camera: bounds on range and range rate accuracy," in *Proc. IEEE Intelligent Vehicles Symposium*, Columbus, OH, July 2003, pp. 120–125.

[10] I. Gat, M. Benady, and A. Shashua, "A monocular vision advance warning system for the automotive aftermarket," in *Proc. SAE World Congress and Exhibition*, Detroit, MI, April 2005, pp. 403–410.

[11] S.H Qi, J Li, Z.P. Sun, J.T. Zhang, and Y. Sun, "Distance estimation of monocular based on vehicle pose information," in *Journal of Physics: Conference Series*, Daqing, China, December 2018, vol. 1168, pp. 1–8.

[12] R. Mur-Artal, J. M. M. Montiel, and J. D Tardos, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, August 2015.

[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, March 2004.

[14] J. Košecká and W. Zhang, "Video compass," in *Proc. European Conference on Computer Vision*, Copenhagen, Denmark, May 2002, pp. 476–490.

[15] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, June 2012, pp. 3354–3361.

[16] StradVision, "https://stradvision.com/," 2021.

A GRAPH CONVOLUTIONAL NEURAL NETWORK FOR RELIABLE GAIT-BASED HUMAN RECOGNITION

Md Shopon, Svetlana Yanushkevich, Yingxu Wang and Marina Gavrilova

University of Calgary, Alberta, Canada

ABSTRACT

In a domain of human-machine autonomous systems, gait recognition provides unique advantages over other biometric modalities. It is an unobtrusive, widely-acceptable way of identity, gesture and activity recognition, with applications to surveillance, border control, risk prediction, military training and cybersecurity. Trustworthy and reliable person identification from videos under challenging conditions, when a subject's walk is occluded by environmental elements, bulky clothing or a viewing angle, is addressed in this paper. It proposes a novel deep learning architecture based on Graph Convolutional Neural Network (GCNN) for accurate and reliable gait recognition from videos. The optimized feature map of the proposed GCNN architecture ensures that recognition remains accurate and invariant to viewing angle, type of clothing or other conditions.

Index Terms— Gait Recognition, Behavioral Biometric, Graph Convolutional Neural Network, Video Processing

1. INTRODUCTION

In a domain of human-machine autonomous systems, an accurate identification of a person from a distance plays an important role. As such, biometric-based gait recognition established itself as one of the best ways to perform a person identification or an activity recognition [1]. It is an unobtrusive, widely-acceptable way of remote identification from videos, that is used in surveillance, border control, medicine, risk prediction, combat training and cybersecurity [2, 3, 4]. Trustworthy and reliable person identification from videos under challenging conditions, when a subject's walk is occluded by environmental elements, bulky clothing, viewing angle or affected by lighting, is a difficult problem [5, 6, 7]. To address it, this paper proposes a novel deep learning architecture based on Graph Convolutional Neural Network (GCNN) for accurate and reliable gait recognition from videos.

Domains of autonomous systems, human-machine interactions, data analytics, and information security traditionally relied on the use of hand-crafted features. However, the rise in cognitive architectures and deep learning methods paved a way to a new avenue for explorations in biometric research [8, 9, 10, 11, 12]. A standard deep learning model for image classification is based on the Convolutional Neural Network (CNN) [13]. It can handle the compositionality of data, local connectivity, and shift-invariance. However, this architecture is unsuitable for gait recognition task under varied conditions, as it lacks ability to be view-invariant and walking direction invariant [13]. Graph Convolutional Neural Networks (GCNN) demonstrated their potential in a variety of applications [14], however they have not been used for biometric gait recognition, due to the challenge of transforming gait sequences to a graph-based representation. The main advantage of GCNN over other architectures is that it allows to explore the spatio-temporal relationship between body joints. Such relationship contains distinguishable features which enhance the performance of gait recognition algorithm.

This work presents the first GCNN based architecture for reliable gait recognition from videos. A novel architecture of graph convolutional neural network is proposed to capture the distinctive spatio-temporal relationships between body joints. We refer to those as kinematic features in the rest of the paper. Next, the proposed architecture utilizes the dynamic modality of the skeleton sequences to achieve the high resistance to low-quality data, varied walking conditions, and different viewing angles. A set of experiments is conducted on CASIA-B gait dataset to confirm that the proposed method outperforms the recent state-of-the-art methods.

2. RELATED WORK

In the last few years, deep learning-based techniques have gained interest in research on video-based gait recognition. In 2014, Wu et al. [15] utilized Convolutional Neural Network (CNN) with hand-crafted histogram and introduced deep features to the network for better performance. In 2017, Liao et al. [16] developed a Pose-Based Temporal-spatial Network (PTSN) architecture to handle specifically carrying and clothing variations and in 2018 He et al. [17] employed Multi-task Generative Adversarial Networks (MGAN) for gait recognition. A good set of distinguishable features can be extracted using deep learning algorithms as long as the hyperparameters are selected properly. However, the computation time for training these deep learning-based methods is high.

Skeleton-based gait recognition approaches have captured a lot of attention recently. Pose estimation algorithms [18]

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

are capable of extracting body joint coordinates of a person from an image or a video. Among the different pose estimation algorithms, OpenPose achieved higher accuracy of joint coordinate prediction [18]. Furthermore, OpenPose algorithm was developed in a way that it gives the flexibility to select a number of body joints to be predicted. Shaik et al. [19] used OpenPose [18] to extract hand-crafted features from the predicted body joints of individual frames. Although this utilized the body joint relationship, the introduced architecture has a large number of trainable parameters. Liao et. al [16] utilized a Pose-based Temporal-Spatial Network (PTSN) to extract the temporal-spatial features to increase recognition accuracy. Although the above-mentioned methods employ the relationship of different body joints between adjacent frames, it is also necessary to propagate information into later frames to accurately recognize a person. Minor differences in the gait information are significant to distinguish individuals. The proposed method leverages GCCN to devise high-level feature map. The proposed architecture is capable of an accurate gait recognition by utilizing the spatio-temporal relationships among the body joints and passing individual body joints information to its neighboring joints. Furthermore, our proposed approach ensures that recognition remains accurate and invariant to viewing angle, type of clothing or other conditions.

3. METHODOLOGY

In this work, a skeleton-based method for gait recognition using Graph Convolutional Neural Network is proposed. GCNN extracts spatio-temporal features from body joints by analyzing the kinematics relation between different body joints. This leads to increased recognition accuracy.

The method for extracting the body joints and gait cycles from videos is discussed in this sub-section. A pre-trained model of OpenPose [18] pose estimation algorithm was used. Two standard methods exist for gait cycle extraction. The first one is mid-stance where the cycle starting phase is when the distance between two feet is smallest. The second one is double support phase, where the distance between both feet is highest. Earlier works [20] demonstrated that double support phase method possesses more relevant distinguishing gait information in comparison with mid-stance method. Thus, the double support phase method is used to extract gait cycles.

In order to fit into GCNN layer, gait cycle data needs to be transformed into a graph first. In this study, the body joints represent a gait cycle. In the graph, body joints as vertices and links between joints are represented as edges. The edges are comprised of two distinct subsets. The first one represents the link between intra-skeleton joints in each individual frame, denoted as $E_S = \{(v_{ti}, v_{tj})|(i,j)\epsilon H\}$, where H denotes the set of the body joints and v_{ti} and v_{tj} represent the connection between i^{th} body joint and j^{th} body joint at time t. The first subset contains the spatial information of the gait sequence. The second subset of edges represents the intra-frame edges, denoted as $E_F = \{(v_{ti}, v_{(t+1)i})|i\epsilon H\}$. Here, all the edges in E_F represents the trajectory over the frame sequence. Temporal information of the skeleton sequence is carried by this subset. One advantage of the proposed graph construction approach is that it preserves the hierarchical representation of the skeleton sequences. The subsets E_F and E_S are stacked together for passing as the input of the GCNN.

A Graph Convolutional Neural Network (GCNN) architecture for gait recognition, proposed in this work is shown in Fig. 1. This GCNN architecture takes two inputs. The first one is the feature vector of all of the body joints, which comprises the predicted coordinate of that particular body joint. The second input is the adjacency matrix, which represents the connection between different body joints. First, the input goes into a GCNN layer followed by a batch normalization layer. The batch normalization layer normalizes input values of each layer's to have a standard deviation of one and a mean output activation of zero. Rectified Linear Unit (ReLU) is used as the activation function of the GCNN layer. Major advantage of ReLU is the reduced likelihood of the gradient to vanish. The constant gradient of ReLU's results in faster learning. The normalized inputs and the adjacency matrix will be passed into subsequent GCNN layer, which later goes into another batch normalization layer. The output of the batch normalization layer is then transformed into a one-dimensional vector which becomes the input of the Multi-Layer Perceptron (MLP) architecture. MLP is a powerful classifier as it can distinguish data that is not linearly separable, or separable by a hyperplane. MLP architecture is comprised of a stack of three layers. Each of the three layer comprises 512, 256, and 128 fully connected nodes respectively. Furthermore, dropout layer is included in between fully connected layer to prevent overfitting. The last layer of the proposed architecture consists of the 74 nodes which represent the number of participants in the training dataset. Softmax activation function is used in the last layer in the dataset to identify the probability of the predicted person. Softmax classifier is very useful because it transforms the scores to a normalized probability distribution which makes it ideal for multi-class classification. In order to calculate the loss between the original and predicted output, categorical cross-entropy loss is used. AMSGrad variant of Adam optimizer [21] is employed to update the weights of the GCNN model using backpropagation algorithm. One key reason for using Adam optimizer is that it is both computationally and memory inexpensive. Furthermore, rescaling of the gradients has no impact during optimization. Adaptive learning rate is utilized in order to achieve better recognition accuracy. During the training, the learning rate is changed according to the validation accuracy and loss. The whole system was developed in Python programming language. Tensorflow and Spectral library was used for developing the architecture of GCNN.

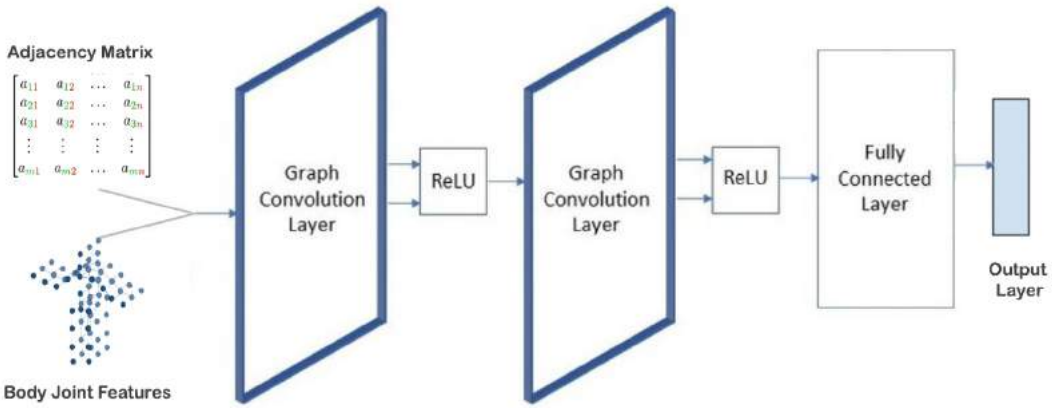


Fig. 1: Architecture of the proposed GCNN model

4. EXPERIMENTATION AND ANALYSIS

CASIA-B is a standard and widely popular dataset for gait recognition and we have chosen this dataset for the experimentation. This dataset includes 124 subjects gait sequence in 3 different walking conditions and all the video sequences were recorded in 11 different viewing angles (0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°). The walking conditions are as follows, 1) Normal Condition (NM) - 6 Sequences, Bag Carrying Condition (BC) - 2 Sequences, 3) Cloth Carrying Condition (CC) - 2 Sequences.

Each subject in the dataset has 110 different video sequences. Data for first 74 subjects are utilized for training, and the rest 50 are kept for testing. From the testing set, the first four normal walking condition videos are kept in gallery (NM 1-4) and last two normal walking conditions (NM 5-6), two cloth carrying (CC 1-2) and two bag carrying (BC 1-2) video sequences are kept in the probe set. The output embedding vector of the 2^{nd} last MLP layer is used for testing purposes. The training set is only used for the adjustment of the model parameters. The gallery set is also known as the query set. The data is put into the model to output the corresponding embedding vector, and then the data of the probe set is also placed in the model to obtain the corresponding embedding vector, and the two vectors are compared. The distance between the embedding vectors of gallery and probe set is obtained using cosine similarity measure. For a probe data whichever gallery data produces the minimum distance is considered to be the class of that probe data.

The proposed method is compared against the two most recent methods specifically developed to handle challenging conditions: the 2017 Pose-based Temporal-Spatial Network (PTSN) [16] and the 2018 Multi-task GAN (MGAN)[17]. The above comparators were chosen because they applied deep learning-based architectures on video-based dataset and showed superior performance to previous methods.

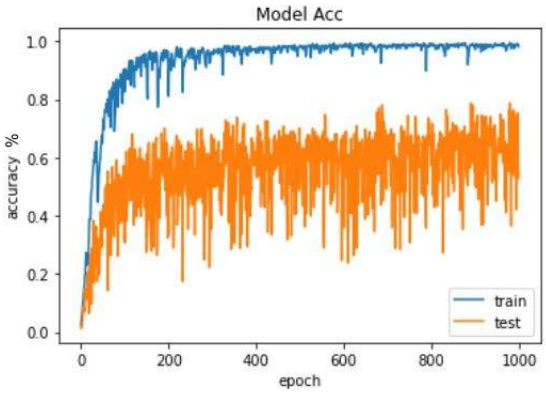
Fig. 2 depicts the learning curves of training and validation set of the proposed method. The learning curve shows the performance of our method over the iteration during training. Moreover, visualizing the learning curve can be a metric to identify the overfitting of the model. From the loss curve we can observe that the loss starts at 6.121 cost and slowly decreased to 1.423 cost. As the loss decreases and accuracy increases over time, we can confirm that the model is not overfitting. However, the model did not generalize the validation set as well as the training set. This can be fixed by incorporating more variety of training data.

CASIA-B dataset consists of videos of three walking conditions and they are normal walking, bag carrying and cloth carrying condition. The proposed method outperforms other state-of-the-art methods for all three walking conditions, as seen from Table 1. For normal walking conditions the proposed method achieved on average 25% higher accuracy over PTSN [16] and 16% over MGAN [17]. Recognition accuracy for the cloth carrying condition of the proposed GCNN is 44% higher than MGAN and 51% higher than PTSN [16]. For the bag carrying condition, the proposed method attained on average 22% higher accuracy than MGAN and 39% higher accuracy than PTSN.

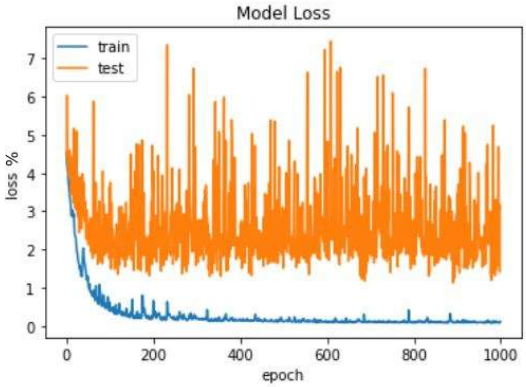
Table 1 also shows a detailed comparison of the proposed architecture for different viewing angles against the compared methods. For bag carrying condition, the method outperforms PTSN on average by 39%, and MGAN on average by 22% with highest gain noted to 82.28% for 54°, 81.42% for 72° and 80.47% for 144° angles. For cloth carrying condition, the method outperforms PTSN on average by 51%, and MGAN on average by 44% with highest gain noted to 80.15% for 72°, 81.11% for 144° and 80.17% 162° angles. Thus, it can be concluded that the proposed method is capable of handling challenging walking conditions. One of the reasons behind achieving a higher accuracy by GCNN is that the proposed method incorporated pose estimation algorithm for predicting the body joints and later utilized joints as the features for the GCNN architecture.

Gallery: Normal Condition#1-4		0° - 180°											Mean
Probe		0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
Normal Condition #5-6	PTSN (Liao et al. 2017)[16]	49.3	61.5	64.4	63.6	63.7	58.1	59.9	66.5	64.8	56.9	44.0	59.3
	MGAN (He et al. 2018) [17]	54.9	65.9	72.1	74.8	71.1	65.7	70.0	75.6	76.2	68.6	53.8	68.1
	Proposed GCNN	**74.23**	**76.52**	**77.32**	**82.15**	**86.78**	**87.92**	**89.81**	**79.32**	**83.25**	**84.82**	**81.88**	**84.32**
Bag Carrying #1-2	PTSN (Liao et al. 2017)[16]	29.8	37.7	39.2	40.5	43.8	37.5	43.0	42.7	36.3	30.6	28.5	37.2
	MGAN (He et al. 2018)[17]	48.5	58.5	59.7	58.0	53.7	49.8	54.0	61.3	59.5	55.9	43.1	54.7
	Proposed GCNN	**67.32**	**72.83**	**74.58**	**82.28**	**81.42**	**74.64**	**72.82**	**76.25**	**80.47**	**78.75**	**79.12**	**76.46**
Cloth Carrying #1-2	PTSN (Liao et al. 2017)[16]	18.7	21.0	25.0	25.1	25.0	26.3	28.7	30.0	23.6	23.4	19.0	24.2
	MGAN (He et al. 2018)[17]	23.1	34.5	36.3	33.3	32.9	32.7	34.2	37.6	33.7	26.7	21.0	31.5
	Proposed GCNN	**65.39**	**72.38**	**74.21**	**76.44**	**80.15**	**74.29**	**73.04**	**72.29**	**81.11**	**80.17**	**78.08**	**75.23**

Table 1: Comparison of rank-1 accuracies of the proposed GCNN against PTSN [16] and MGAN [17] on CASIA-B dataset for Normal, Bag Carrying and Cloth Carrying conditions.



(a) Training vs validation accuracy



(b) Training vs validation loss

Fig. 2: Accuracy and loss graph of the proposed method

5. CONCLUSION

In this work, we propose a Graph Convolutional Neural Network based method for gait recognition. The proposed Graph Convolutional Neural Network method aggregates the extracted features from one frame to another by utilizing the kinematics relationship between body joints. The performance analysis on the publicly accessible CASIA-B dataset shows that the proposed approach is superior to the current state-of-the-art methods. Our method outperforms all of the compared research works on the CASIA-B dataset for all of the conditions, with the highest improvements attained for the cloth carrying condition. If we look at the results for all conditions and view angles, our method shows consistent performance, proving that the proposed method is view-invariant. In the future, we will investigate how to improve GCNN architecture by incorporating residual connection. In addition, various pooling methods will be investigated and real-time implementation for surveillance will be explored.

6. ACKNOWLEDGEMENT

The authors would like to acknowledge the National Sciences and Engineering Research Council of Canada for partial support of this research in the form of the NSERC Discovery Grant 10007544 and the IDEaS network funding AutoDefence: Towards Trustworthy Technologies for Autonomous Human-Machine System.

7. REFERENCES

[1] Qinghan Xiao, "Technology review biometrics technology, application, challenge, and computational intelligence solutions," *IEEE Computational Intelligence Magazine*, vol. 2, no. 2, pp. 5–25, 2007.

[2] Mohammad S Obaidat, P Venkata Krishna, V Saritha, and Shubham Agarwal, "Advances in key stroke dynamics-based security schemes," in *Biometric-Based Physical and Cybersecurity Systems*, pp. 165–187. Springer, 2019.

[3] Ann-Kathrin Seifert, Abdelhak M Zoubir, and Moe-

ness G Amin, "Radar-based human gait recognition in cane-assisted walks," in *2017 IEEE Radar Conference*. IEEE, 2017, pp. 1428–1433.

[4] Chien-Wen Cho, Wen-Hung Chao, Sheng-Huang Lin, and You-Yin Chen, "A vision-based analysis system for gait recognition in patients with parkinson's disease," *Expert Systems with Applications*, vol. 36, no. 3, pp. 7033–7039, 2009.

[5] Svetlana N Yanushkevich and Mark S Nixon, *Image pattern recognition: synthesis and analysis in biometrics*, vol. 67, World Scientific, 2007.

[6] Yingxu Wang, Svetlana Yanushkevich, Ming Hou, Konstantinos Plataniotis, Mark Coates, Marina Gavrilova, Yaoping Hu, Fakhri Karray, Henry Leung, and Arash Mohammadi, "A tripartite theory of trustworthiness for autonomous systems," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020, pp. 3375–3380.

[7] Marina L Gavrilova, Ferdous Ahmed, ASM Hossain Bari, Ruixuan Liu, Tiantian Liu, Yann Maret, Brandon Kawah Sieu, and Tanuja Sudhakar, "Multi-modal motion-capture-based biometric systems for emergency response and patient rehabilitation," in *Research Anthology on Rehabilitation Practices and Therapy*, pp. 653–678. 2021.

[8] Yingxu Wang, Bernard Widrow, Lotfi A Zadeh, Newton Howard, Sally Wood, Virendrakumar C Bhavsar, Gerhard Budin, Christine Chan, Rodolfo A Fiorini, and Marina L Gavrilova, "Cognitive intelligence: deep learning, thinking, and reasoning by brain-inspired systems," *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, vol. 10, no. 4, pp. 1–20, 2016.

[9] Yingxu Wang, Newton Howard, Janusz Kacprzyk, Ophir Frieder, Phillip Sheu, Rodolfo A Fiorini, Marina L Gavrilova, Shushma Patel, Jun Peng, and Bernard Widrow, "Cognitive informatics: Towards cognitive machine learning and autonomous knowledge manipulation," *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, vol. 12, no. 1, pp. 1–13, 2018.

[10] Yingxu Wang, Mehrdad Valipour, Omar D Zatarain, Marina L Gavrilova, Amir Hussain, Newton Howard, and Shushma Patel, "Formal ontology generation by deep machine learning," in *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*. IEEE, 2017, pp. 6–15.

[11] Stanley Tarng, Deng Wang, and Yaoping Hu, "Estimating cognitive processes related to haptic interaction within virtual environments," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2019, pp. 2823–2828.

[12] Yingxu Wang, Fakhri Karray, Sam Kwong, Konstantinos N Plataniotis, Henry Leung, Ming Hou, Edward Tunstel, Imre J Rudas, Ljiljana Trajkovic, Okyay Kaynak, et al., "Perspectives on the philosophical, cognitive and mathematical foundations of symbiotic autonomous systems (sas)," *Philosophical Transactions of Royal Society (A), Oxford, UK, 379(X):1-20*, 2021.

[13] Yann LeCun and Yoshua Bengio, "Convolutional networks for images, speech, and time series," *The Handbook of Brain Theory and Neural Networks*, vol. 3361, no. 10, pp. 1995, 1995.

[14] Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski, "Graph convolutional networks: a comprehensive review," *Computational Social Networks*, vol. 6, no. 1, pp. 1–23, 2019.

[15] Zifeng Wu, Yongzhen Huang, Liang Wang, Xiaogang Wang, and Tieniu Tan, "A comprehensive study on cross-view gait based human identification with deep cnns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 2, pp. 209–226, 2016.

[16] Rijun Liao, Chunshui Cao, Edel B Garcia, Shiqi Yu, and Yongzhen Huang, "Pose-based temporal-spatial network (PTSN) for gait recognition with carrying and clothing variations," in *Chinese Conference on Biometric Recognition*. Springer, 2017, pp. 474–483.

[17] Yiwei He, Junping Zhang, Hongming Shan, and Liang Wang, "Multi-task gans for view-specific feature learning in gait recognition," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 102–113, 2018.

[18] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh, "Openpose: realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2019.

[19] Shahil Shaik, *OpenPose based Gait Recognition using Triplet Loss Architecture*, Ph.D. thesis, Dublin, National College of Ireland, 2020.

[20] Rawesak Tanawongsuwan and Aaron Bobick, "Modelling the effects of walking speed on appearance-based gait recognition," in *CVPR 2004.* IEEE, 2004, vol. 2, pp. 1–8.

[21] Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 2015.

MULTI-SCALE FEATURE FUSION: LEARNING BETTER SEMANTIC SEGMENTATION FOR ROAD POTHOLE DETECTION

Jiahe Fan[1]*, Mohammud J. Bocus*[2]*, Brett Hosking*[3]*, Rigen Wu*[4]*, Yanan Liu*[2]*, Sergey Vityazev*[5]*, Rui Fan*[6⋆]

[1]Beijing Institute of Technology, Beijing 100811, P. R. China.
[2]University of Bristol, Bristol, BS8 1TL, United Kingdom.
[3]Arm, Manchester, M1 3HU, United Kingdom.
[4]ATG Robotics, Hangzhou 310000, P. R. China.
[5]Ryazan State Radio Engineering University, Ryazan 390005, Russia.
[6]Tongji University, Shanghai 201804, P. R. China.
Email: jhxfan@ieee.org, junaid.bocus@bris.ac.uk, brett.hosking@arm.com, wrg6370@outlook.com, yl17692@bris.ac.uk, vityazev.s.v@ieee.org, rui.fan@ieee.org

ABSTRACT

This paper presents a novel pothole detection approach based on single-modal semantic segmentation. It first extracts visual features from input images using a convolutional neural network. A channel attention module then reweighs the channel features to enhance the consistency of different feature maps. Subsequently, we employ an atrous spatial pyramid pooling module (comprising of atrous convolutions in series, with progressive rates of dilation) to integrate the spatial context information. This helps better distinguish between potholes and undamaged road areas. Finally, the feature maps in the adjacent layers are fused using our proposed multi-scale feature fusion module. This further reduces the semantic gap between different feature channel layers. Extensive experiments were carried out on the Pothole-600 dataset to demonstrate the effectiveness of our proposed method. The quantitative comparisons suggest that our method achieves the state-of-the-art (SoTA) performance on both RGB images and transformed disparity images, outperforming three SoTA single-modal semantic segmentation networks.

Index Terms— pothole detection, single-modal semantic segmentation, convolutional neural network, feature fusion.

1. INTRODUCTION

Potholes are considerable structural failures on the road surface [1]. They are caused by the contraction and expansion of the road surface as rainwater permeates the ground [2]. The affected road areas are further deteriorated due to tire vibration. This makes the road surface impracticable for driving [3]. The vehicular traffic can cause subsurface materials to move, which further expands the potholes, creating a vicious circle [4]. To avoid traffic accidents, it is crucial and necessary to detect road potholes in time [5]. With recent advances in machine learning, automated road pothole detection systems have become a reality [6–9]. Benefiting from the evolution of convolutional neural networks (CNNs), semantic segmentation has become an effective technique for road pothole detection [5], and it has achieved compelling results.

Among the state-of-the-art (SoTA) semantic segmentation CNNs, fully convolutional network (FCN) [10] replaces the fully connected layer used in traditional classification networks with a convolutional layer to achieve better segmentation results. Contextual information aggregation has proved to be an effective tool that can be used to improve segmentation accuracy. ParseNet [11] captures global context by concatenating global pooling features. PSPNet [12] introduces a spatial pyramid pooling (SPP) module to collect contextual information in different scales. Atrous SPP (ASPP) [13–15] applies different dilated convolutions to capture multi-scale contextual information without introducing extra parameters.

To take advantage of global contextual visual information, some pioneering methods have been proposed to reweigh 2-D feature map channels. SE-Net [16] and EncNet [17] are designed to learn a globally-shared attention vector from global context. SE-Net [16] employs a squeeze-excitation operation to integrate the global contextual information into a feature weight vector and reweigh the feature maps. EcnNet [17] uses a context encoding module to obtain a globally-shared feature weight vector. This module adopts learning and residual encoding components to obtain a global context encoded feature vector, which is then used to predict the feature weight vector. Combining global context information to reweigh the feature map of each channel has proved to be effective in terms of

⋆ Corresponding Author

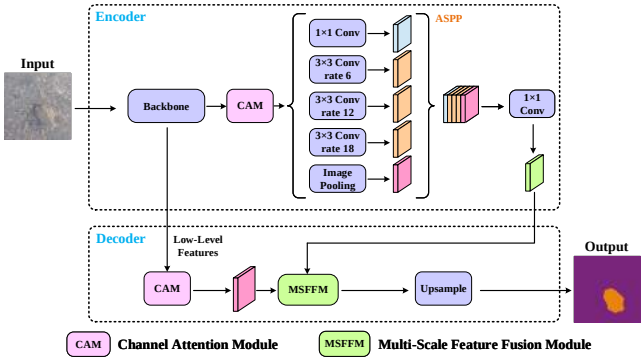


Fig. 1. The architecture of our proposed road pothole detection network.

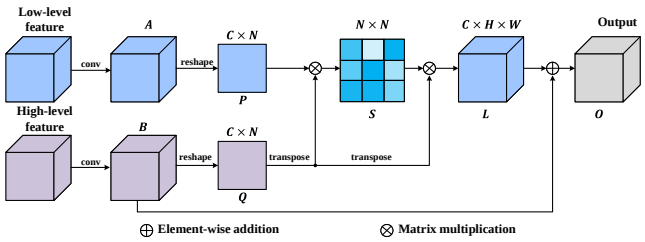


Fig. 2. Our proposed Multi-Scale Feature Fusion Module.

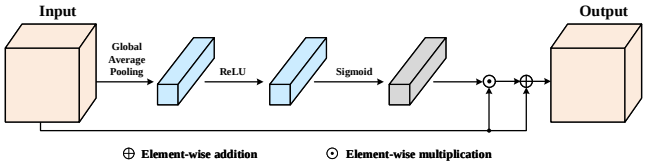


Fig. 3. Our employed Channel Attention Module.

improving semantic segmentation accuracy.

Some other methods use backbone CNNs [12, 14, 17, 18] to extract feature maps at different scales. By performing a series of convolution and pooling operations, the top layer has rich semantic information [19–22], while the lower-level feature maps contain fine-grained information [23]. This information asymmetry becomes a barrier to accurate semantic prediction. To address this issue, U-Net [24] adopts an encoder-decoder architecture to improve the semantic segmentation performance. It adds skip connections between the encoder and decoder, which can recover fine-grained details in the semantic prediction. Feature pyramid network (FPN) [25] uses the structure of U-Net [24] with predictions from each level of the feature pyramid. However, the fusion operations cannot measure the semantic relevance between feature maps at different scales. The semantic information between feature maps at different scales may interfere with each other.

To address the above problems, in this paper, we propose a novel multi-scale feature fusion module (MSFFM) based on attention mechanism. Our main objective is to improve the semantic prediction by leveraging additional low-level information near the boundaries, where the pixel categories are difficult to infer. We utilize a matrix multiplication operation to measure the relevance between the two feature maps in the spatial dimension, which is the basic idea of weight vectors. By reweighing feature maps in lower layers, we reduce interference between feature maps in different layers. Moreover, we adopt a channel attention module (CAM) to reweigh feature maps in different channels to further improve the semantic segmentation results.

2. METHODOLOGY

Given a road image, potholes can have diverse shapes and scales. We can obtain feature maps at the top layer through a series of convolution and pooling operations. Although the feature maps have rich semantic information, their resolutions are not high enough to provide accurate semantic prediction. Unfortunately, directly combining low-level feature maps can only bring very limited improvements. To overcome this shortcoming, we design an effective feature fusion module in this paper.

The schema of our proposed road pothole detection network is illustrated in Fig. 1. Firstly, we employ a pre-trained dilated ResNet-101 as the backbone CNN to extract visual features. We also replace the down-sampling operations with dilated convolutions in the last two ResNet-101 [26] blocks, thus the size of the final feature map is 1/8 of the input image. This module helps retain more details without introducing extra parameters. In addition, we adopt the ASPP module used in Deeplabv3 [14] to collect contextual information in the top feature map. Then, we adopt a CAM to reweigh the feature maps in different channels. It can highlight some features so as to produce better semantic predictions. Finally, we feed the feature maps at different levels into the MSFFM to improve the segmentation performance near the pothole contour.

2.1. Multi-scale feature fusion

The top feature maps have rich semantic information but their resolution is low, especially near the pothole boundary. On the other hand, the lower feature maps have low-level semantic information but higher resolution. In order to address this problem, some works [15, 24, 27] directly combine the feature maps in different layers. Nevertheless, their achieved improvements are very limited because of the semantic gap between feature maps with different scales.

The attention modules have been widely applied in many works [28–30]. Inspired by some successfully applied spatial attention mechanisms, we introduce a MSFFM, which is based on spatial attention to efficiently fuse the feature maps

at different scales. Semantic gap is one of the key challenges in feature fusion. To solve this issue, the MSFFM calculates the correlation between pixels in different feature maps via matrix multiplication, and the correlation is then utilized as the weight vectors for the higher-level feature map:

$$s_{ji} = \frac{\exp(P_i \cdot Q_j)}{\sum_{i=1}^{N} \exp(P_i \cdot Q_j)}, \quad (1)$$

where s_{ji} measures the relevance between the $i\text{-}th$ position in lower feature map and the $j\text{-}th$ position in higher feature map. N represents the number of pixels. P and Q represent the lower and higher feature maps generated by the convolutional layer, respectively, where $\{P, Q\} \in \mathbb{R}^{C\times N}$. The higher the similarity between feature representations of pixels at the two positions, the greater is the relevance between them. As shown in Fig. 2, we first feed the feature maps into a convolution layer to compress the channels for fewer calculations while generating feature maps A and B, $\{A, B\} \in \mathbb{R}^{C\times H\times W}$. H and W represent the height and width of the feature map. Then we reshape the low-level feature map A and the high-level feature map B to P and Q, respectively, where $N = H \times W$ represents the number of pixels. Afterwards, we transpose Q for matrix multiplication and apply a softmax layer to calculate the spatial attention map $S \in \mathbb{R}^{N\times N}$.

Then we perform matrix multiplication between Q and the spatial attention map S to generate the feature map $L \in \mathbb{R}^{C\times H\times W}$. Finally, we utilize an element-wise sum operation between B and L to obtain the final output $O \in \mathbb{R}^{C\times H\times W}$ as follows:

$$O_j = \alpha \sum_{i=1}^{N} (s_{ji} \mathrm{q}_i) + B_j, \quad (2)$$

where α is initialized as 0 and it gradually learns to assign more weight, q_i represents the $i\text{-}th$ position in the lower feature map, and B_j represents the $j\text{-}th$ channel of the top feature map. It can be inferred from (2) that each position of the final feature O is a weighted sum of the features across all positions of the top features. As the final feature is generated by the top features, the high-level semantic information is well preserved in the final outputs.

In summary, we utilize matrix multiplication to measure the relevance of pixels in feature maps from different layers, which integrates the detailed information from the lower feature map into the final outputs, thus improving the semantic segmentation performance for the pothole boundary. We apply this module between the last two layers.

2.2. Channel-wise feature reweighing

It is well-known that high-level features have rich semantic information and each channel map can be regarded as a class-specific response. Each response can affect the final semantic prediction to a different extent. Therefore, we utilize CAMs, as illustrated in Fig. 3, to enhance the consistency of the feature maps in each layer, by changing the features' weights in each channel. The CAM is designed to reweigh each channel according to the overall pixels of each feature map. We first employ a global average pooling layer to squeeze spatial information. Subsequently, we use the Rectified Linear Unit (ReLU) and sigmoid function to generate the weight vectors, which are finally combined with the input feature maps by element-wise multiplication operations to generate an output feature map. The overall information is integrated into the weight vectors, making the feature maps more reliable and the pothole detection results closer to the ground truth. In our experiments, we employ the CAM in the 4th and 5th layers.

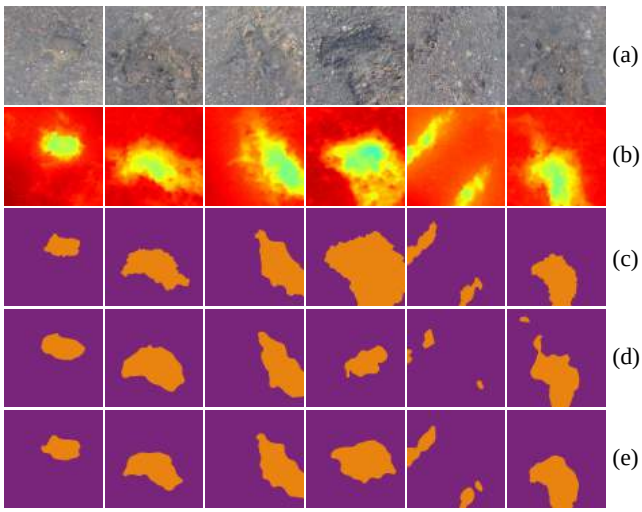


Fig. 4. Examples of pothole detection results: (a) RGB images; (b) transformed disparity images; (c) pothole ground truth; (d) semantic RGB image segmentation results; (e) semantic transformed disparity image segmentation results.

3. EXPERIMENT RESULTS

In this paper, we carry out comprehensive experiments on the Pothole-600 dataset [4] to evaluate the performance of our proposed road pothole detection both qualitatively and quantitatively. This dataset provides two modalities of vision sensor data: 1) RGB images, and 2) transformed disparity images [31]. The transformed disparity images were obtained by performing disparity transformation [32, 33] on dense disparity images estimated by PT-SRP [34]. We conduct experiments to select the best architecture. All the experiments use the same training setups.

Ablation Study: To validate the effectiveness of our proposed MSFFM and CAM, we first carry out the ablation study on different network architectures, as shown in Table 1 and

Table 1. Ablation study on RGB images.

Methods	mIoU (%)	mFsc (%)
Baseline	55.32	71.23
Baseline + CAM	57.17	72.75
Baseline + MSFFM	59.43	74.55
Baseline + CAM + MSFFM (ours)	61.51	76.16

Table 2. Ablation study on transformed disparity images.

Methods	mIoU (%)	mFsc (%)
Baseline	70.90	82.97
Baseline + CAM	72.26	83.89
Baseline + MSFFM	71.02	83.06
Baseline + CAM + MSFFM (ours)	72.75	84.22

Table 3. Performance of other SoTA networks on RGB images.

Methods	mIoU (%)	mFsc (%)
PSPNet [12]	58.61	73.90
DANet [18]	59.42	74.54
Deeplabv3 [15]	58.60	73.90

Table 4. Performance of other SoTA networks on transformed disparity images.

Methods	mIoU (%)	mFsc (%)
PSPNet [12]	69.85	82.25
DANet [18]	70.52	82.71
Deeplabv3 [15]	70.36	82.60

Table 2. The baseline network uses Deeplabv3 [14], which concatenates the feature maps from ASPP module and the lower layer.

Moreover, we implement the two modules into the baseline network and verify their effectiveness, respectively. According to the results shown in Table 1 and Table 2, implementing two modules can achieve better performance than the baseline network on both RGB images and transformed disparity images. The mIoU improvements on RGB images with the use of CAM and MSFFM are 1.85% and 4.11%, respectively, while the mIoU improvements on the transformed disparity images are 1.36% and 0.12%, respectively. The network with MSFFM and CAM embedded yields an mFsc of 76.16% on RGB images and an mFsc of 84.22% on transformed disparity images. Based on these experimental results, we believe that the CAM and MSFFM adopted in our network can improve the segmentation accuracy significantly.

Performance Comparison: We also compare our method with three SoTA semantic segmentation CNNs: 1) Deeplabv3 [15], 2) PSPNet [12], 3) DANet [18] on both RGB images and transformed disparity images, as shown in Table 3 and Table 4. PSPNet [12] and Deeplabv3 [15] collect contextual information in different scales, and therefore, they achieve similar results on RGB images and transformed disparity images. DANet [18] collects contextual information based on attention mechanism and it shows better performance on both RGB images and transformed disparity images. This further demonstrates the superiority of attention mechanism on semantic segmentation for road pothole detection, which can also be observed from the comparison between our method and other SoTA networks.

Additionally, when using RGB images, the mIoUs of our method are 2.91%, 2.9%, and 2.09% higher than those achieved by Deeplabv3 [15], PSPNet [12], and DANet [18], respectively. Moreover, our method also outperforms the above-mentioned SoTA semantic segmentation networks on transformed disparity images, where the improvements on mIoU with respect to Deeplabv3 [15], PSPNet [12], and DANet [18] are 2.39%, 2.9%, and 2.23%, respectively. Specifically, our method achieves the best performance, even when it only utilizes a MSFFM.

We also provide some qualitative results of our proposed road pothole detection method in Fig. 4, where it can be observed that the CNN achieves accurate results on the transformed disparity images. The results obtained from our comprehensive experimental evaluations have demonstrated the effectiveness and superiority of our method compared to other SoTA techniques. Owing to the proposed CAM and MSFFM, our method achieves better performance for potholes detection on both RGB and transformed disparity images.

4. CONCLUSION

This paper introduced a method to detect road potholes based on semantic segmentation, which employs a novel multi-scale feature fusion module based on spatial attention to reduce the semantic gap between the feature maps in different layers. This helps maintain the semantic information in the higher-level feature maps and combine the detailed information near the pothole boundary. The top feature maps can be reweighed using the vectors generated by the relevance of each pixel in the different layers, which combine the global information of the feature maps. Moreover, a channel attention module is introduced to strengthen the channels which are more relevant to the semantic segmentation ground truth. Extensive experiments were conducted on both RGB images and transformed disparity images, where our proposed network outperforms all other SoTA semantic segmentation networks.

5. REFERENCES

[1] Rui Fan et al., "Pothole detection based on disparity transformation and road surface modeling," *IEEE Transactions on Image Processing*, vol. 29, pp. 897–908, 2019.

[2] John S Miller et al., "Distress identification manual for the long-term pavement performance program," Tech. Rep., 2003.

[3] Senthan Mathavan et al., "A review of three-dimensional imaging technologies for pavement distress detection and measurements," *IEEE TITS*, 2015.

[4] Rui Fan et al., "We learn better road pothole detection: from attention aggregation to adversarial domain adaptation," in *European Conference on Computer Vision*. Springer, 2020, pp. 285–300.

[5] Rui Fan et al., "Rethinking road surface 3d reconstruction and pothole detection: From perspective transformation to disparity map segmentation," *IEEE Transactions on Cybernectics*, 2021.

[6] Hengli Wang et al., "Applying surface normal information in drivable area and road anomaly detection for ground mobile robots," *IROS*, 2020.

[7] Rui Fan et al., "Road crack detection using deep convolutional neural network and adaptive thresholding," in *2019 IEEE Intelligent Vehicles Symposium*. IEEE, 2019.

[8] Christian Koch and Ioannis Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, 2011.

[9] Rui Fan et al., "Sne-roadseg: Incorporating surface normal information into semantic segmentation for accurate freespace detection," in *European Conference on Computer Vision*. Springer, 2020, pp. 340–356.

[10] Jonathan Long et al., "Fully convolutional networks for semantic segmentation," in *CVPR*, 2015.

[11] Wei Liu et al., "Parsenet: Looking wider to see better," *CoRR*, 2015.

[12] Hengshuang Zhao et al., "Pyramid scene parsing network," in *CVPR*, 2017, pp. 2881–2890.

[13] Liang-Chieh Chen et al., "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE TPAMI*, 2017.

[14] Liang-Chieh Chen et al., "Rethinking atrous convolution for semantic image segmentation," *CoRR*, 2017.

[15] Liang-Chieh Chen et al., "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *ECCV*, 2018, pp. 801–818.

[16] Jie Hu, Li Shen, and Gang Sun, "Squeeze-and-excitation networks," in *CVPR*, 2018, pp. 7132–7141.

[17] Hang Zhang et al., "Context encoding for semantic segmentation," in *CVPR*, 2018, pp. 7151–7160.

[18] Jun Fu et al., "Dual attention network for scene segmentation," in *CVPR*, 2019.

[19] Liang-Chieh Chen et al., "Semantic image segmentation with deep convolutional nets and fully connected crfs," *CoRR*, 2014.

[20] David Eigen and Rob Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in *CVPR*, 2015.

[21] Fayao Liu et al., "Deep convolutional neural fields for depth estimation from a single image," in *CVPR*, 2015, pp. 5162–5170.

[22] Fayao Liu et al., "Learning depth from single monocular images using deep convolutional neural fields," *IEEE TPAMI*, 2015.

[23] Guosheng Lin et al., "Refinenet: Multi-path refinement networks for high-resolution semantic segmentation," in *CVPR*, 2017, pp. 1925–1934.

[24] Olaf Ronneberger et al., "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*. Springer, 2015.

[25] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie, "Feature pyramid networks for object detection," in *CVPR*, 2017.

[26] Kaiming He et al., "Deep residual learning for image recognition," in *CVPR*, 2016, pp. 770–778.

[27] Vijay Badrinarayanan et al., "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE TPAMI*, vol. 39, no. 12, pp. 2481–2495, 2017.

[28] Zhouhan Lin and et al., "A structured self-attentive sentence embedding," *CoRR*, 2017.

[29] Ashish Vaswani et al., "Attention is all you need," in *NeurIPS*, 2017.

[30] Tao Shen et al., "Disan: Directional self-attention network for rnn/cnn-free language understanding," in *AAAI*, 2018, vol. 32.

[31] Rui Fan et al., "Real-time dense stereo embedded in a uav for road inspection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019.

[32] Rui Fan and Ming Liu, "Road damage detection based on unsupervised disparity map segmentation," *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[33] Hengli Wang et al., "Dynamic fusion module evolves drivable area and road anomaly detection: A benchmark and algorithms," *IEEE Transactions on Cybernectics*, 2021.

[34] Rui Fan et al., "Road surface 3d reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3025–3035, 2018.

DEEP LEARNING ARCHITECTURES USED IN EEG-BASED ESTIMATION OF COGNITIVE WORKLOAD: A REVIEW

Nusrat Zerin Zenia, Yaoping Hu

Department of Electrical and Software Engineering
University of Calgary
Calgary, AB, CANADA

ABSTRACT

Cognitive workload (CWL) refers to the ratio of a participant's mental effort over his/her brain capacity when executing tasks with aid of a machine. Such CWL influences the participant's trust placed on the machine and thus affects the tasks' performance. Efficient human-machine interaction demands the machine's real-time adaptation to meet an admissible CWL for the participant. The adaptation needs estimating CWL based on brain activities captured by non-invasive electroencephalography (EEG). Since deep learning (DL) is common for extracting EEG features reflecting certain characteristics of the activities, DL-based CWL estimation attracts ample attention. Herein, we present a review to summarize current trends in DL architectures for EEG-based CWL estimation and to identify gaps in the trends for future work.

***Index Terms*—** Cognitive workload, EEG, deep learning

1. INTRODUCTION

Cognitive workload (CWL) is defined as a quantitative measure of mental efforts forced on cognitive resources (e.g., working memory) of the human brain while performing a task [1]. As the brain resources are constrained, tasks overloading cognition may reduce efficiency and result in critical errors. In contrast, a deficient workload is a waste of the resources and leads to boredom during task execution [2]. In addition, CWL is inversely proportional to trust – an intrinsic characteristic needed for an autonomous human-machine system (HMS) [3, 4]. An HMS requires collaboration between a participant and a machine to bring forth flexible decision-making [5]. The collaboration would ideally be constructive if the participant could trust the machine to undertake designated actions. The trust needs the machine to be adaptive for ensuring the participant's admissible CWL. Estimating the CWL is thus crucial to assure the performance of the HMS.

Two approaches are prevalent in estimating CWL – subjective and objective measures [6]. Subjective measures – such as NASA task load index (NASA-TLX) [7], Bedford-scale [8], etc. – allow a participant to answer questionnaires by using certain self-rating while (or after) performing a task. The self-rating is easy to be administrated but potentially influenced by the participant's honesty (i.e., subjectivity). In contrast, objective measures use behavioral and physiological recordings logged during the participant's execution of a main task. The behavioral recordings involve conventionally a secondary task, which causes changes of the participant's CWL [9]. Without disrupting the main task, the physiological recordings – like electroencephalography (EEG), heart rate, eye movement, etc. – are promising to estimate the CWL. Among these recordings, non-invasive EEG is most common because of its propensity for placing electrodes on the scalp to log brain activities as signals [10].

Many approaches of EEG-based CWL estimation use some types of machine learning [11]. A major drawback of these types is their need of handcrafted attributes for computation. Such attributes may disregard salient information embedded in high-dimensional EEG signals (as inputs). In turn, CWL estimation may be inaccurate or erroneous. A remedy is deep learning (DL). Being able to extract high-level features from inputs, a DL architecture with more than two hidden layers is commonly used in many efforts [10] – such as classifying EEG features related to mental states [12, 13] and extracting event-related potentials from EEG inputs for estimating CWL [14]. But an unexplored topic is the trend and gaps of DL architectures for estimating CWL.

Herein, this paper provides a systematic review on using DL for EEG-based CWL estimation. The main goal of this review is three folds: to provide a methodological guideline for formulating EEG inputs, to report the trend and gaps of existing DL architectures for estimating CWL, and to recommend directions for future research.

2. METHODS

2.1. Search strategy

We identified relevant studies via a literature search. Keywords used in the search were "EEG", "cognitive workload or mental workload", and "deep learning". The search spanned over 5 years from 2016 to 2021, covering the majority of work. Electronic databases involved in the search were Sci-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

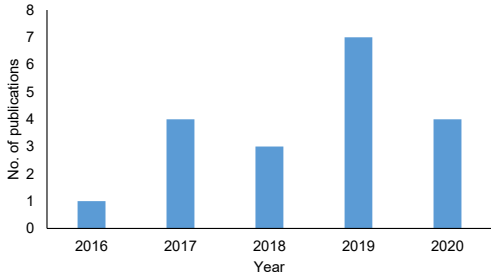


Fig. 1: Temporal distribution of all selected studies.

enceDirect, ACM Digital Library, IEEE Xplore, and Google Scholar. The final date of the search was on 18 February 2021.

2.2. Inclusion and exclusion criteria

We applied inclusion and exclusion criteria to select the identified studies for their pertinence. The inclusion criteria accounted that the studies estimated CWL based only on EEG, used DL, and were in English. Due to the page limit, the exclusion criteria considered that the studies were not in English and/or published as reviews, books, and theses.

2.3. Selection of studies

Across the databases, our keyword searching retrieved 1380 records. Title/abstract screening, along with duplicate removal, yielded 50 studies for further consideration. The inclusion and exclusion criteria determined finally 19 studies pertinent for the review, as depicted in Fig. 1.

3. RESULTS AND DISCUSSIONS

We scrutinized the 19 studies under various attributes including task, EEG format, features, EEG preprocessing, DL architecture, hyper-parameter, and accuracy. Table 1 summarizes the outcomes of the scrutinizing.

3.1. Task and workload level

To measure CWL, a participant needs to execute a task inducing various difficulties of cognition. We defined the induced difficulties as workload levels (WL). That is, WL is a measure of the participant's CWL when executing the task. As shown in Table 1, WL equals 2 representing binary levels of high and low; 3 being high, medium, and low; and so on. All studies under review used relative WL, without providing an absolute value to constitute a level.

To simulate the task, many studies used a software application called automation-enhanced cabin air management system (ACAMS). The application ACAMS instructed a participant to perform a safety-critical task to control the air quality of a cabin [29, 6, 28, 27]. Similarly, other studies used various simulators mimicking scenarios of driving a vehicle or flying an aircraft for measuring CWL [20, 19, 17]. Another common task, namely an N-back task, required the participant to mentally recall a stimulus to match the Nth letter prior to the stimulus [30, 2, 10]. The task of mental arithmetic was also used frequently for measuring CWL [1, 18, 25].

3.2. EEG processing

A non-invasive EEG employs electrodes placed on the scalp to capture brain activities as voltage signals. Varying in time and space, the signals are highly sensitive to undesired eye and muscle movements (i.e., noise) [12]. Thus, raw EEG signals need to be preprocessed before inputting to a DL architecture. Among the studies under review, common preprocessing techniques were fast Fourier transform (FFT), downsampling, filtering, and independent component analysis (ICA). The majority of the studies applied more than one technique for preprocessing. As depicted in Fig. 2(a), filtering is the most pervasive technique adopted by 11 of the 19 studies. Downsampling and FFT rank in the second, and four studies use ICA. Other techniques including normalization, segmentation, and data augmentation are occasionally applied for EEG preprocessing.

After preprocessing, EEG signals need to be transformed into a particular format before inputting to a DL architecture. Among the studies under review, three common formats were image, extracted features (EF), and signal. The image format includes 2D pictures, spectral maps, topographic maps, recurrence plots, etc. Figure 2(b) shows that 37% of the studies use the image format, another 37% apply the EF format, and 21% utilize the signal format. Interestingly, only 5% of the studies employ raw EEG signals as inputs without preprocessing.

Many EF could be outputs of a DL architecture and then inputs of another architecture in the EF format. Spectral features (SF) were dominant among the studies under review. Common SF used in the studies were power spectral density (PSD). Convolutional neural network (CNN) and recurrent neural network (RNN) – two common variants of DL architecture – often extracted temporal features (TF) and spatial features (SpF). The combination of TF and SpF as inputs enhanced the performance accuracy of other architectures. Figure 3(a) plots EEG input formats versus DL architectures used in the studies. As depicted in Fig. 3(a), CNN mainly takes the image format, whereas RNN prefers the signal and EF formats. Variants of CNN exploit all three formats.

3.3. DL architecture

The choice of a DL architecture impacts on the correctness and speed of CWL estimation. Figure 3(b) depicts the percentage of the studies under review using different DL architectures. The most prevalent architecture is CNN (26%), followed by the combination of CNN and RNN (21%). Variants of CNN – like ensemble CNN and multiple stream CNN – account for about 16% of the studies. However, the popularity of CNN was contingent based on its ability to extract SpF

Table 1: Summary of the studies under review

Ref.	Participants$^{\alpha}$	Task (WL)$^{\beta}$	EEG format$^{\gamma}$	Features (#channel)$^{\theta}$	Preprocessing$^{\Gamma}$	DL Architecture$^{\zeta}$	Hyper-parameter$^{\eta}$	Accuracy$^{\phi}$
[1]	15, F-8, M-7	Memor. (4)	Image	SF, TF, SpF (3)	FFT	CNN+RNN	#HL-7, #BNTM-3, Ac-ReLU, G-D(0.5)	96.3%
[15]	48	SIMKAP (3)	Signal	SF, TF (2)	Ft	BLSTM	#HL-9, G-D(0.2), N-Batch	82.57%
[16]	18, F-2, M-16	Navigation (2)	RSC	RSC (64)	Ds, Ft, ICA	CNN	#HL-4, Ac-Softmax, G-D(0.5)	93%
[17]	18, F-1, M-6	Simulated flight(3)	Signal	SF, TF, SpF (30)	Ds, Ft, ICA, Sg	MFB-CNN	#HL-5, Ac-ReLU, G-D(0.5), N-Batch	75%
[2]	17, F-0, M-17	N-back tasks (3)	TM	TF, SF (16)	Ft, D. aug.	CNN+TCN	#HL-7, #TCN-7, Ac-ReLU, softmax, G-D(0.2), N-Batch	91.9%
[10]	20, F-0, M-20	N-back and arithmatic (2)	TM	SF, TF, SpF (16)	Ft, D. aug., Ds	CNN+RNN	#HL-7, Ac-sigmoid, G-D(0.5), N-Batch	88.9%
[18]	13	Memor. (4)	SFE	SF	FFT	CNN	VGG-19, N-Batch	93.71%
[19]	16	Dn simul. (4)	RP	SF (3)	Ft, Ds	CNN	Incepion-V3	≈87.6%
[20]	1, M-1	Driving simulator (3)	Raw EEG	Not specified (4)	Sl, Z-score Norm	CNN	#HL-8, Ac-ReLU, G-D(0.5), N-Batch	96%
[21]	8	ACAMS (2)	EF	PSD,TF (11)	Ft, IC, FFT	EL-SDAE	#HL-2, Ac-Sigmoid	92%
[22]	46	ACAMS (2)	EF	PSD,TF (11)	Ds, Sg	TDAE	#HL-4, Ac-sigmoid	90%
[23]	25, F-10	Not specified (4)	Image	SF, TF, SpF	Ft, FFT	CNN+RNN	#HL-9, #LSTM-2, Ac-ReLU, Sigmoid, Softmax; G-D(0.5);	92.5%
[24]	8, F-2, M-6	MATB (2)	Signal	SF, TF	Ds, FFT	CNN+RNN	#HL-6, #BLSTM-1, Ac-Sigmoid	≈ 86%
[25]	15, F-8, M-7	Memor. (4)	Image	SF, TF (21)	Ds, FFT	CNN	#HL- 4, Ac - ReLU	92.37%
[26]	7	ACAMS (2)	EF	PSD (11)	Ft, FFT, ICA	SDAE	#HL-6, Ac-Sigmoid	95.5%
[27]	6, F-0, M-6	ACAMS (4)	EF	SF	Ft, STFT	ECNN	Not specified	93.8%
[28]	6, F-0, M-6	ACAMS (4)	EF	MAD	Ft, Ds	RBM	#HL-2, Ac-Softmax	≈96.1%
[5]	6	MATB (3)	EF	SF, TF (19)	Not specified	RNN	#HL-2, Ac-Sigmoid, G-D(0.2)	93 %
[6]	8, F-0, M-8	ACAMS (2)	EF	PSD (11)	ICA, FFT	SDAE	#HL-5, Ac-sigmoid	74%

[Note: α: M - Male, F - Female

β: ACAMS - automatic enhanced cabin air management system, SIMKAP - simultaneous capacity based multitasking activity, MATB - multi attribute task battery, Memor. - Memorization, Dr. simul. - Drone simulator

γ: RSC - retrosplenial complex, SFE - spectral feature enhanced map, EF - extracted features

θ: SF - spectral feature, TF - temporal feature, SpF - spatial feature, PSD - power spectral density

ζ: CNN - convolutional neural network, RNN - recurrent neural network, MFB - multiple feature block, TCN - temporal convolutional network, SDAE - stacked denoising autoencoder, EL-SDAE - ensemble SDAE, ECNN - ensemble CNN, TDAE - transfer dynamical autoencoder, RBM - restricted Boltzmann machine

Γ: HL- hidden layer, Ac - activation function, G - generalization, D(*) - dropout (rate), Op - optimization, N - normalization

η: ICA - independent component analysis , FFT - fast Fourier transform, STFT - short-time Fourier transform, RP - recurrent plot, MAD - mean absolute deviation, Ft - Filtering, Ds - Downsampling, Sl - Slicing, Sg - Segmentation, D. aug. - Data augmentation]

and TF automatically from EEG signals without handcrafted attributes. Complimenting CNN, RNN learns temporal dependencies among other features over a large time interval. The learning eventually aided in improving the performance accuracy of both RNN and CNN together. Another DL architecture, namely stacked denoising auto encoder (SDAE), comprised 16% of the studies and was mainly for alleviating cross-participant and cross-feature variations. This alleviation paved a way to explore using the process of transfer learning for measuring CWL.

For a DL architecture, its hyper-parameters determine its structure such as the number of hidden layers, activation functions, dropout rates, and use of batch normalization [13]. The structure plays an important role in achieving computational performance (i.e., resources and time). Therefore, it is crucial to set up the pertinent values of the hyper-parameters. As indicated in Table 1, the studies under review have the number of hidden layers from 2 to 10, and use widely a rectified linear or a Sigmoid activation function. In addition, the studies apply commonly a dropout rate of 0.2 and batch normalization for generalization and input uniformity, respectively.

The accuracy of DL architectures reflects the correctness

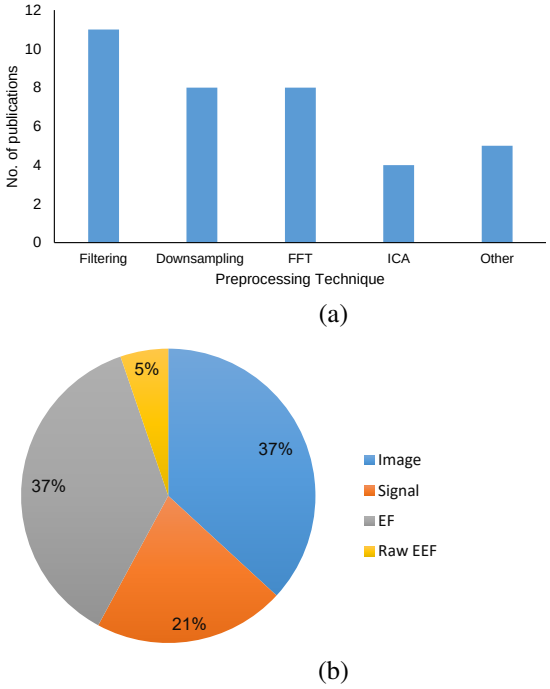


Fig. 2: EEG processing techniques: (a) distribution of the techniques among the studies under review and (b) distribution of input formats among the studies.

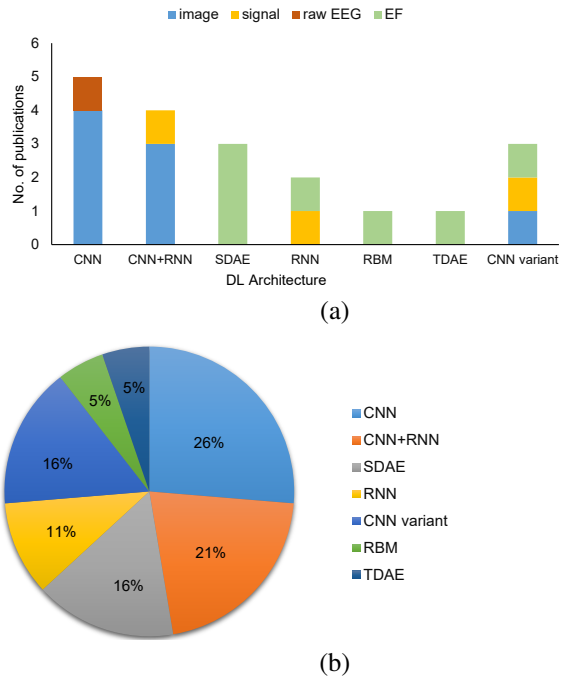


Fig. 3: DL architectures in the studies under review: (a) relationship between DL architectures and their input formats and (b) percentage of the studies using these architectures.

of classifying estimated CWL into different levels. The majority of the studies claim more than 90% accuracy of their classification, as shown in Table 1. The studies use however various experimental setups and nonuniform interpretations of the levels. It is thus not pragmatic to compare achieved accuracy among the studies. Moreover, all of the studies reveal no computational time to obtain the accuracy. This might arise from an observation that DL architectures are in general computationally very expensive [31].

3.4. Discussions

Although both EEG and DL play notable roles in investigating brain activities, there are few efforts to correlate subjective measures of CWL and EEG features [32]. Besides, the studies under review use various tasks and interpret the levels of CWL differently. Some studies utilize a within-participant design for feature classification, and other studies employ a cross-participant design for the very classification. It is thus impossible to infer a DL architecture to be most suitable for estimating CWL. Furthermore, the studies focus mainly on improving the accuracy of DL architectures but ignore a gap caused by their long computation time. With a growing need for facilitating real-time interactivity between a human and a machine, a shorter time is crucial to enable the machine adapting to the human's CWL accordingly. This shorter time could equal to a cognitive cycle of the human brain, which sustains a duration of about 400 ms [33]. This duration sets an upper limit for completing CWL estimation, providing feedback derived from the estimation to the machine, and activating the machine to make an adaptive change. Currently, few research efforts are devoted to addressing this gap for achieving fast CWL estimation. Future investigations are thus imperative to fill the gap through multiple potential pathways.

One pathway could use machine learning (ML) for CWL estimation. If handcrafted attributes are preexisted, the computation of ML is faster than that of DL. Being time consuming, handcrafting attributes need to be completed offline ahead of implementing ML for CWL estimation. Another pathway might be optimization of DL architectures to trade off between accuracy and computation time.

4. CONCLUSION

A significant impact of CWL is on human trust within HMS. Our review suggested a potential of combining EEG and DL for CWL estimation, albeit with certain DL shortcomings. Future work addressing the shortcomings would be necessary.

5. ACKNOWLEDGEMENT

The authors acknowledge that this review is supported by the Department of National Defence's Innovation for Defence Excellence and Security (IDEaS) program, Canada and by an NSERC Alliance – Alberta Innovates Advance grant, Canada.

6. REFERENCES

[1] W. Qiao and X. Bi, "Ternary-task convolutional bidirectional neural turing machine for assessment of EEG-based cognitive workload," *Biomed. Signal Process. Control.*, vol. 57, pp. 101745, Mar. 2020.

[2] P. Zhang et al., "Spectral and temporal feature learning with two-stream neural networks for mental workload assessment," *IEEE Trans. Neural Syst*, vol. 27, no. 6, pp. 1149–1159, 2019.

[3] K. Gupta et al., "In ai we trust: investigating the relationship between biosignals, trust and cognitive load in vr," in *Proc. ACM VRST*, New York, NY, USA, 2019, pp. 1–10.

[4] M. Novitzky et al., "Preliminary interactions of human-robot trust, cognitive load, and robot intelligence levels in a competitive game," in *ACM/IEEE HRI*, Chicago, IL, USA, 2018, p. 203–204.

[5] R. G. Hefron et al., "Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation," *Pattern Recognition Lett.*, vol. 94, pp. 96–104, 2017.

[6] Z. Yin and J. Zhang, "Recognition of cognitive task load levels using single channel eeg and stacked denoising autoencoder," in *Proc. IEEE CCC*, Chengdu, China, 2016, pp. 3907–3912.

[7] S. G.Hart and L. E. Staveland, "Development of nasa-tlx (task load index): results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, 1988.

[8] A. H. Roscoe and G. A. Ellis, "A subjective rating scale for assessing pilot workload in flight: a decade of practical use," Tech. Rep., Royal Aerospace Establishment Farnborough (UK), Mar. 1990.

[9] K. Moustafa et al., "Assessment of mental workload: a comparison of machine learning methods and subjective assessment techniques," in *Proc. H-Workload*, Dublin, Ireland, 2017, pp. 30–50.

[10] P. Zhang et al., "Learning spatial–spectral–temporal eeg features with recurrent 3d convolutional neural networks for cross-task mental workload assessment," *IEEE Trans. Neural Syst*, vol. 27, no. 1, pp. 31–42, 2019.

[11] M. Plechawska-Wojcik et al., "A three-class classification of cognitive workload based on eeg spectral data," *Applied Sciences*, vol. 9, pp. 5340, 2019.

[12] Alexander Craik, Yongtian He, and Jose L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *J. Neural Eng.*, vol. 16, no. 3, pp. 031001, Apr. 2019.

[13] Y. Roy et al., "Deep learning-based electroencephalography analysis: a systematic review," *J. Neural Eng.*, vol. 16, no. 5, pp. 051001, Aug. 2019.

[14] U. Ghani et al., "ERP based measures of cognitive workload: A review," *Neurosci. Biobehav. Rev.*, vol. 118, pp. 18–26, 2020.

[15] D. Chakladar et al., "EEG-based mental workload estimation using deep BLSTM-LSTM network and evolutionary algorithm," *Biomed. Signal Process. Control.*, vol. 60, pp. 101989, July 2020.

[16] T. T. N. Do et al., "Estimating the cognitive load in physical spatial navigation," in *Proc. IEEE SSCI*, Canberra, ACT, Australia, 2020, pp. 568–575.

[17] D. Lee et al., "Continuous eeg decoding of pilots' mental states using multiple feature block-based convolutional neural network," *IEEE Access*, vol. 8, pp. 121929–121941, 2020.

[18] Y. Zhang and Y. Shen, "Parallel mechanism of spectral feature-enhanced maps in eeg-based cognitive workload classification," *Sensors*, vol. 19, no. 4, 2019.

[19] Q. Zhang et al., "Identifying mental workload using eeg and deep learning," in *Proc. IEEE CAC*, Hangzhou, China, 2019, pp. 1138–1142.

[20] M. A. Almogbel et al., "Cognitive workload detection from raw eeg-signals of vehicle driver using deep learning," in *Proc. ICACT*, PyeongChang, Korea (South), 2019, pp. 1–6.

[21] S. Yang et al., "Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders," *Comput. Biol. Med.*, vol. 109, pp. 159–170, 2019.

[22] Z. Yin et al., "Physiological-signal-based mental workload estimation via transfer dynamical autoencoders in a deep learning framework," *Neurocomputing*, vol. 347, pp. 212–229, June 2019.

[23] S. Kuanar et al., "Cognitive analysis of working memory load from eeg, by a deep recurrent neural network," in *Proc. IEEE ICASSP*, Calgary, AB, Canada, 2018, pp. 2576–2580.

[24] R. Hefron et al., "Cross-participant EEG-based assessment of cognitive workload using multi-path convolutional recurrent neural networks," *Sensors*, vol. 18, no. 5, Apr. 2018.

[25] Z. Jiao et al., "Deep convolutional neural networks for mental load classification based on eeg data," *Pattern Recognit.*, vol. 76, pp. 582–595, 2018.

[26] Y. Zhong and Z. Jianhua, "Cross-session classification of mental workload levels using eeg and an adaptive deep learning model," *Biomed. Signal Process. Control.*, vol. 33, pp. 30–47, 2017.

[27] J. Zhang et al., "Pattern classification of instantaneous mental workload using ensemble of convolutional neural networks," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14896–14901, 2017.

[28] J. Zhang and S. Li, "A deep learning scheme for mental workload classification based on restricted Boltzmann machines," *Cogn. Technol. Work*, vol. 19, no. 4, pp. 607–631, Nov. 2017.

[29] J. Sauer et al., "Designing automation for complex work environments under different levels of stress," *Appl. Ergon.*, vol. 44(1), pp. 119–27, 2013.

[30] C. Herff et al., "Mental workload during n-back task quantified in the prefrontal cortex using fnirs," *Front. Hum. Neurosci.*, vol. 7, pp. 935, 2014.

[31] S. Shi et al., "Benchmarking state-of-the-art deep learning software tools," in *IEEE CCBD*, Macau, China, 2016, pp. 99–104.

[32] G. Funke et al., "Evaluation of subjective and EEG-based measures of mental workload," in *Proc. HCI*, Berlin, Heidelberg, 2013, pp. 412–416.

[33] T. Madl et al., "The timing of the cognitive cycle," *PLoS ONE*, vol. 6, no. 4, Apr. 2011.

SIMULTANEOUS DISTRIBUTED ESTIMATION AND ATTACK DETECTION/ISOLATION IN SOCIAL NETWORKS: STRUCTURAL OBSERVABILITY, KRONECKER-PRODUCT NETWORK, AND CHI-SQUARE DETECTOR

Mohammadreza Doostmohammadian[†⋆], Themistoklis Charalambous[†], Senior Member, IEEE, Miadreza Shafie-khah[*], Senior Member, IEEE, Nader Meskin[◇], Senior Member, IEEE, and Usman A. Khan[‡], Senior Member, IEEE

† School of Electrical Engineering, Aalto University, Espoo, Finland.
⋆ Faculty of Mechanical Engineering, Semnan University, Semnan, Iran.
* School of Technology and Innovations, University of Vaasa, Vaasa, Finland
◇ Electrical Engineering Department, Qatar University, Doha, Qatar
‡Electrical and Computer Engineering Department, Tufts University, Medford, MA, USA.

ABSTRACT

This paper considers distributed estimation of linear systems when the state observations are corrupted with Gaussian noise of unbounded support and under possible random adversarial attacks. We consider sensors equipped with single time-scale estimators and local chi-square (χ^2) detectors to simultaneously observe the states, share information, fuse the noise/attack-corrupted data locally, and detect possible anomalies in their own observations. While this scheme is applicable to a wide variety of systems associated with full-rank (invertible) matrices, we discuss it within the context of distributed inference in social networks. The proposed technique outperforms existing results in the sense that: (i) we consider Gaussian noise with no simplifying upper-bound assumption on the support; (ii) all existing χ^2-based techniques are centralized while our proposed technique is distributed, where the sensors *locally* detect attacks, with no central coordinator, using specific probabilistic thresholds; and (iii) no local-observability assumption at a sensor is made, which makes our method feasible for large-scale social networks. Moreover, we consider a Linear Matrix Inequalities (LMI) approach to design block-diagonal gain (estimator) matrices under appropriate constraints for isolating the attacks.

Index Terms— Attack detection and isolation, Kronecker-product network, distributed estimation, χ^2-test.

1. INTRODUCTION

The unprecedented large size of social networks mandates distributed sensing, inference, and detection [1–7], where the information is collected and processed locally while meeting certain security concerns. Recent distributed estimation protocols [6–11] are prone to faults/attacks that may result in inaccurate state estimates. Different attack detection and FDI (fault detection and isolation) strategies are thus proposed in the literature, ranging in applications from biological modeling [12] to smart-grid monitoring [13, 14], and from centralized approaches [15–20] to more recent distributed methods [21–23]. Among the centralized solutions, deterministic FDI and attack detection methods design decision thresholds based on the upper-bound on the noise support [17, 18], while, in contrast, probabilistic χ^2-test with no such assumption on the noise is proposed in [15] and further developed in [19, 20]. Among the distributed strategies, [23] requires injecting a watermarking input signal conceding to a loss in the control/estimation performance, which is not applicable to *autonomous* systems (such as the social network model in this paper). In order to close this gap, this paper aims at developing a technique for distributed inference of autonomous (social) systems while simultaneously detecting and isolating adversarial attacks *locally* with no central coordinator.

The main contributions of this paper are as follows. (i) This work considers a windowed χ^2 benchmark to *locally* design probabilistic decision thresholds based on certain false alarm rates (FARs). This is in contrast to existing *deterministic* thresholds assuming certain upper-bound on the noise support [17, 18], which results in faulty outcome when the noise upper-bound is considerably larger than the attack/fault magnitude. (ii) This work extends the recent *centralized* χ^2 detectors [19, 20] to *distributed* ones, where the sensors are widespread over a large social network and, thus, the centralized solutions are infeasible/undesirable due to heavy communication loads or inability for parallel processing. In this direction, the notion of *Kronecker-product network* [24] is used to perceive (structural) observability of the composite social/sensor network, which allows to find

This work has been supported by the European Commission through the H2020 project FinEst Twins under grant agreement No 856602. The work of U. Khan was supported by NSF under awards #1903972 and #1935555. The work of T. Charalambous was supported by the Academy of Finland under Grant 317726. Corresponding email: doost@semnan.ac.ir, mohammadreza.doostmohammadian@aalto.fi.

minimal connectivity requirement on the sensor network for distributed estimation/detection. (iii) Our distributed technique, as in [21, 22], does not require local-observability at every sensor. However, unlike *fixed* biasing faults/attacks on sensor outputs in [21, 22], this work extends the results to general anomalies in the form of a *random* variable. In particular, we adopt the notion of *distance measure* [19], a scalar variable to compare the residual variance in presence and absence of attacks.

2. PROBLEM FORMULATION

We consider the interaction of individuals in a social network as a linear-structure-invariant (LSI) autonomous model [4–7],

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \nu_k, \qquad k \geq 0, \tag{1}$$

where k is the time-step, A is the social system matrix associated with social digraph $\mathcal{G}$, $\nu_k \sim \mathcal{N}(0, Q)$ is additive i.i.d noise vector, and vector $\mathbf{x}_k = \left[x_k^1, \ldots, x_k^n\right]^\top \in \mathbb{R}^n$ represents the global social state. Note that n is the size of social network and $\mathbf{x}_k^i$ represents the i'th individual's social state, e.g., opinion, rumor, or attitude [1–7]. The state $\mathbf{x}_k^i$ of individual i at time k is a weighted average of the states $\mathbf{x}_{k-1}^j$ of its in-neighbors in $\mathcal{G}$ and its *own previous state* $\mathbf{x}_{k-1}^i$. This is well-justified for opinion-dynamics in social systems, and particularly implies that matrix A is (structurally) full-rank [7]. Consider N social sensors (agents or information-gatherers [5]) sensing the state of some individuals as,

$$y_k^i = H_i\mathbf{x}_k + \tau_k^i + \eta_k^i, \tag{2}$$

with H_i as the measurement matrix, τ_k^i as possible attack and $\eta_k^i \sim \mathcal{N}(0, R_i)$ as Gaussian noise at sensor i at time k. Define $R = \text{diag}[R_i]$ as the covariance matrix of the i.i.d noise vector η_k. Throughout this paper, without loss of generality, we assume every sensor observes one state variable, i.e., $y_k^i \in \mathbb{R}$. Further, as in similar works [15, 19, 20], we assume the system and measurement noise covariance (Q and R) are known. Then, sensors share their information over a sensor network $\mathcal{G}_N$. Clearly, system A is not locally observable to any sensor, but globally observable to all sensors. The condition on (A, H)-observability is given in the following lemma.

Lemma 1 *[7] Given a social network $\mathcal{G}$ (with structurally full-rank adjacency matrix A), if at least one social state is sensed in every strongly-connected-component (SCC) in $\mathcal{G}$, then, the pair (A, H) is (structurally) observable.*

Given (social) system (1) and state observations (2) satisfying Lemma 1, we aim to design a distributed iterative procedure to simultaneously estimate the (social) state $\mathbf{x}_k^i$ while detecting adversary attacks at (social) sensors. The attack by the adversary is modeled as an additive random term τ_k^i at sensor i in (2). The proposed distributed estimation makes the entire system observable to every sensor, and the attack-detection technique enables each sensor to locally detect anomalies in its observation with a certain FAR (false-alarm rate).

3. MAIN ALGORITHM

We consider a modified version of the single time-scale distributed estimator in [7] with one step of averaging on *a-priori* estimates (similar to *DeGroot consensus model* [8]) and one step of measurement update (also known as *innovation* [8]),

$$\widehat{\mathbf{x}}_{k|k-1}^i = \sum_{j\in\mathcal{N}(i)} w_{ij} A\widehat{\mathbf{x}}_{k-1|k-1}^j, \tag{3}$$

$$\widehat{\mathbf{x}}_{k|k}^i = \widehat{\mathbf{x}}_{k|k-1}^i + K_i H_i^\top \left(y_k^i - H_i\widehat{\mathbf{x}}_{k|k-1}^i\right). \tag{4}$$

with *stochastic matrix* $W = \{w_{ij}\}$ as the adjacency matrix of the sensor network $\mathcal{G}_N$ representing the fusion weights among the sensors, K_i as the local gain matrix at agent i, and $\widehat{\mathbf{x}}_{k|k-1}^i$ and $\widehat{\mathbf{x}}_{k|k}^i$ as the state estimate at time k given all the information of agent i and its in-neighbors $\mathcal{N}(i)$, respectively, at time $k-1$ and k. In contrast to double time-scale estimators/observers [11] with many consensus iterations between every two consecutive time-steps $k-1$ and k of social dynamics (1), the estimator (3)-(4) performs one iteration of information fusion between steps $k-1$ and k, which is more efficient in terms of computation/communication loads.

Define the estimation error at agent i as $\mathbf{e}_k^i = \mathbf{x}_k - \widehat{\mathbf{x}}_{k|k}^i$ and the error vector $\mathbf{e}_k = \left[\mathbf{e}_k^{1\top}, \ldots, \mathbf{e}_k^{n\top}\right]^\top$. Following similar procedure as in [6], the error dynamics is as follows,

$$\mathbf{e}_k = (W \otimes A - KD_H(W \otimes A))\mathbf{e}_{k-1} + \mathbf{q}_k, \tag{5}$$

with $D_H = \text{diag}[H_i^\top H_i]$, $K = \text{diag}[K_i]$ as the feedback gain matrix, and $\mathbf{q}_k$ as the collective vector of noise-related terms $\mathbf{q}_k = \left[\mathbf{q}_k^{1\top}, \ldots, \mathbf{q}_k^{n\top}\right]^\top$ as,

$$\mathbf{q}_k^i = \nu_{k-1} - K_i\Big(H_i^\top\eta_k^i + H_i^\top\tau_k^i + H_i^\top H_i\nu_{k-1}\Big), \tag{6}$$

$$\mathbf{q}_k = \mathbf{1}_N \otimes \nu_{k-1} - KD_H(\mathbf{1}_N \otimes \nu_{k-1}) - K\overline{D}_H\eta_k - K\overline{D}_H\tau_k, \tag{7}$$

with $\mathbf{1}_N$ as the vector of 1's of size N and $\overline{D}_H = \text{diag}[H_i^\top]$. Following Kalman theory, for bounded steady-state estimation error, $(W \otimes A, D_H)$ needs to be observable, characterizing the *distributed observability* condition for network of estimators/observers [25]. Using structured system theory, this condition can be investigated via graph theoretic notions. In this direction, the associated network to $W \otimes A$ is a Kronecker-product network, whose observability condition relies on the structure of both $\mathcal{G}$ and $\mathcal{G}_N$. Given the social network $\mathcal{G}$, the conditions on the sensor network $\mathcal{G}_N$ to satisfy distributed observability follows the recent results on *composite-network theory* and network observability discussed in [24], which is summarized in the following lemma.

Lemma 2 *[24] Given (A, H)-observability via Lemma 1, minimal sufficient condition for $(W \otimes A, D_H)$-observability is that matrix W be irreducible, i.e., the network $\mathcal{G}_N$ be strongly-connected (SC).*

For an observable pair $(W \otimes A, D_H)$, the feedback gain matrix K can be designed to stabilize the error dynamics (5). Mathematically, for $\overline{A} = W \otimes A - KD_H(W \otimes A)$, we need to design K such that $\rho(\overline{A}) < 1$ (Schur stability of error dynamics (5)) for general social systems with $\rho(A) > 1$ with $\rho(\cdot)$ as the spectral radius. As mentioned before, for distributed case, K needs to be further block-diagonal such that each sensor only uses local information in its own neighborhood. The iterative LMI-based algorithm to design such block-diagonal gain K is given in [26]. In attack-free scenario, the distributed estimator/observer (3)-(4) with proper gain K ensures tracking the global social state with bounded steady-state error as discussed in [6, 7]. Next, in this section, we further study the performance of the proposed protocol in the presence of non-zero random attack signals. Define $\widehat{y}_k^i = H_i \widehat{\mathbf{x}}_{k|k}^i$ as the estimated output at sensor i at time k. To detect possible attacks, each sensor calculates its *residual* as the difference of its original output and the estimated one,

$$r_k^i = y_k^i - \widehat{y}_k^i = y_k^i - H_i \widehat{\mathbf{x}}_{k|k}^i = H_i \mathbf{e}_k^i + \eta_k^i + \tau_k^i. \quad (8)$$

Having $\rho(\overline{A}) < 1$, the steady-state error in (5) only relies on the term $\mathbf{q}_k^i$ defined in (6) as,

$$\begin{aligned} H_i \mathbf{q}_k^i = {} & H_i \nu_{k-1} - H_i K_i H_i^\top \eta_k^i \\ & - H_i K_i H_i^\top \tau_k^i - H_i K_i H_i^\top H_i \nu_{k-1}. \end{aligned} \quad (9)$$

From (8) and (9), it is clear that for $\tau_k^i \neq 0$, only the residual r_k^i at sensor i is biased with no effect on the residual of other sensors $j \neq i$. This allows to *isolate* the attacked sensor as r_k^i only depends on τ_k^i and not on τ_k^j's. To detect a possible attack at agent i via residual r_k^i, the non-zero term $\tau_k^i - H_i K_i H_i^\top \tau_k^i$ needs to be sufficiently larger than other noise terms. This ensures that the residual in attacked case is large enough to be distinguished from the noise terms in attack-free case. Clearly, the detecting probability of an attack depends on the magnitude of τ_k^i, which justifies the *probabilistic threshold design*. In this direction, we consider a distributed probability-based χ^2-test which outperforms the deterministic fault/attack detection methods as it considers noise of *unbounded support*. In this case, instead of a deterministic threshold with 0 (no attack) or 1 (attack detected) outcome, different probabilistic thresholds (with different sensitivities) are defined each assigned with an FAR. In fact, higher *residual-to-noise ratio* (RNR) stimulates the threshold with lower FAR. In this direction, first the covariance of error $\mathbf{e}_k$ and (attack-free) residuals need to be calculated, which are tied with the noise covariance Q and R. Let $\Xi_k = \mathbb{E}(\mathbf{e}_k \mathbf{e}_k^\top)$ and $\Sigma = \mathbb{E}(\mathbf{q}_k \mathbf{q}_k^\top)$. Then, from (5),

$$\Xi_k = \overline{A}^k \Xi_0 (\overline{A}^k)^\top + \sum_{j=1}^{k-1} \overline{A}^j \Sigma (\overline{A}^j)^\top + \Sigma. \quad (10)$$

Knowing that $\rho(\overline{A}) < 1$, the first term in (10) goes to zero. Therefore, it can be proved from [4] that for $\Xi_\infty = \lim_{k\to\infty} \Xi_k$,

$$\|\Xi_\infty\|_2 = \|\sum_{j=1}^{\infty} \overline{A}^j \Sigma (\overline{A}^j)^\top + \Sigma\|_2 \leq \frac{\|\Sigma\|_2}{1 - b^2}, \quad (11)$$

with $b = \|\overline{A}\|_2 < 1$. For attack-free case ($\tau_k = \mathbf{0}_N$ in (6)),

$$\begin{aligned} \mathbf{q}_k \mathbf{q}_k^\top = {} & (I_{Nn} - KD_H)(\mathbf{1}_{NN} \otimes \nu_{k-1}\nu_{k-1}^\top)(I_{Nn} - KD_H)^\top \\ & + (K\overline{D}_C)\eta_k \eta_k^\top (K\overline{D}_H)^\top, \end{aligned} \quad (12)$$

where $\mathbf{1}_{NN}$ is the 1's matrix of size N. Applying the $\mathbb{E}(\cdot)$ and 2-norm operators,

$$\begin{aligned} \|\Sigma\|_2 = {} & \|(I_{Nn} - KD_H)(\mathbf{1}_{NN} \otimes Q)(I_{Nn} - KD_H)^\top\|_2 \\ & + \|(K\overline{D}_H) R (K\overline{D}_H)^\top\|_2. \end{aligned} \quad (13)$$

Then, the upper-bound on $\|\Sigma\|_2$ is,

$$\|\Sigma\|_2 \leq \|I_{Nn} - KD_H\|_2^2 N \|Q\|_2 + \|K\|_2^2 \|\overline{R}\|_2,$$

with $\overline{R} = \text{diag}[H_i^\top R_i H_i]$. Then, using (11),

$$\frac{\|\Xi_\infty\|_2}{N} \leq \frac{a_1 N \|Q\|_2 + a_2 a_3 \|R\|_2}{N(1 - b^2)} = \Phi, \quad (14)$$

where $\|I_{Nn} - KD_H\|_2^2 = a_1$, $\|K\|_2^2 = a_2$, and $\|\overline{R}\|_2 = a_3 \|R\|_2$. Note that in (14) the error covariance is scaled by the number of sensors N. From (14), assuming no attack is present ($\tau_k^i = 0$), a conservative approximation for error variance at sensor i is $\mathbb{E}(\mathbf{e}_k^i \mathbf{e}_k^{i\top}) = \Phi$. Then, following the discussion in [19], the residual r_k^i in (8) can be assumed as a zero-mean Gaussian variable with maximum variance $\Lambda_i = \mathbb{E}(r_k^i r_k^{i\top}) = H_i^\top \Phi H_i + R_i$, i.e., $r_k^i \sim \mathcal{N}(0, \Lambda_i)$. Define,

$$z_k^i = \frac{(r_k^i)^2}{\Lambda_i}, \; v_k^i = \sum_{t=k-T+1}^{k} z_t^i, \quad (15)$$

with T as the length of the *sliding window*[1]. It is known that, for a Gaussian variable r_k^i, scalars z_k^i and v_k^i follow χ_1^2-distribution with degree 1 and T respectively ($\mathbb{E}[z_k^i] = 1$ and $\mathbb{E}[v_k^i] = T$) [27]. In fact, these so-called *distance measures* z_k^i and v_k^i give an estimate of variance of r_k^i relative to the attack-free variance Λ_i [19], and are known to outperform simple detectors comparing *absolute residual* to a threshold as in [21, 22]. Next, we determine the *decision threshold* θ on v_k^i based on a pre-specified FAR p. It can be shown that

[1] In general, each agent can consider a different length for the horizon T.

$p = 1 - F(\theta)$ where $F(\cdot)$ is the cumulative distribution function (CDF) of χ^2_1-distribution. Then,

$$\theta = 2\Gamma^{-1}(1-p, \frac{T}{2}), \tag{16}$$

with $\Gamma^{-1}(\cdot,\cdot)$ as the *inverse regularized lower incomplete gamma function* [27]. Using (16), our attack detection logic at each sensor i is as follows,

$$\text{If } \begin{cases} v_k^i \geq \theta \\ v_k^i < \theta \end{cases} \text{ Then } \begin{cases} \mathcal{H}_1^i : \text{Attack Detected} \\ \mathcal{H}_0^i : \text{No Attack} \end{cases} \tag{17}$$

It should be noted that the existing χ^2-based attack detection scenarios in literature are all centralized [15, 19, 20] and in this work, using distributed estimation, we enable detection of attacks *locally* at every sensor with no need of a central unit. We summarize our proposed simultaneous distributed estimation and attack detection technique in Algorithm 1.

Algorithm 1: Proposed iterative methodology.

1 **Given:** System matrix A, Network $\mathcal{G}_N$, Fusion matrix W, Measurements y_k, Measurement matrix H, System/Measurement noise covariance Q/R, false-alarm probability (FAR) p, sliding window T
2 Choose block-diagonal gain K via LMI in [26];
3 Find $\widehat{\mathbf{x}}^i_{k|k}$ at every sensor i via (3)-(4);
4 Find Λ_i based on R, Q, and (14);
5 Find residuals r^i_k at every sensor i via (8);
6 Find z^i_k and v^i_k at every sensor i via (15) ;
7 Define threshold θ based on p and T via (16);
8 If $v^i_k \geq \theta$ return $\mathcal{H}^i_1$: Attack Detected with FAR p;
9 If $v^i_k < \theta$ return $\mathcal{H}^i_0$: No Attack;
10 **Return:** Hypothesis $\mathcal{H}^i_0$ or $\mathcal{H}^i_1$ for $i = \{1, ..., N\}$.

Note that after detecting a malicious attack with low FAR, the strategy in [28] can be adopted to remove unreliable data and replace the compromised sensor with its *observationally equivalent* counterpart to regain distributed observability.

4. SIMULATION RESULTS

We evaluate our theoretical results on an example social network $\mathcal{G}$ of 10 state nodes with 4 sensor observations shown in Fig. 1. The network $\mathcal{G}_N$ of 4 sensors is considered as a cycle (satisfying Lemma 2). The fixed non-zero entries of A and W are chosen randomly in $(0, 1.1]$. Further, $\rho(A) = 1.1$ implying a potentially unstable system, $\eta^i_k, \nu^i_k \sim \mathcal{N}(0, 0.06)$, and non-zero entries of H are set as 1. Using MATLAB CVX, the stabilizing block-diagonal gain K is designed via the iterative LMI in [26] subject to $|1 - H_i K_i H_i^\top| > 0.2$, which results in $\rho(\overline{A}) = 0.97$, $b = 1.42$, $\Phi = 4.82$, and $\Lambda_i = 4.88$. In attack-free case, each sensor is able to track the global social state $\mathbf{x}_k$ over time via protocol (3)-(4). The time-evolution

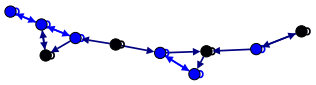


Fig. 1: The small social network $\mathcal{G}$ considered for simulation. The black state nodes are observed by the sensors (satisfying Lemma 1).

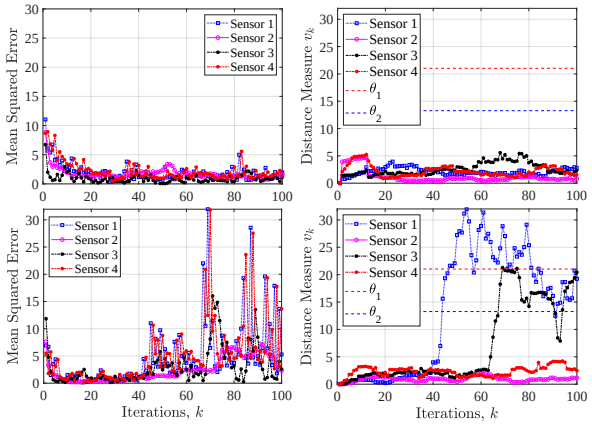


Fig. 2: (top) No attack: mean squared estimation error at all sensors are steady-state stable. (bottom) Attack at sensors 1 and 3: the non-zero attacks add bias to the estimation error at all sensors. Distance measures v^1_k and v^3_k exceeding θ_2 reveal possible attacks at sensors 1 and 3 with FAR $p_2 = 35\%$, while v^1_k exceeding θ_1 implies lower FAR $p_1 = 5\%$ for attack at sensor 1.

of mean squared errors $\frac{\|\mathbf{e}^i_k\|^2}{n}$ and distance measures at all sensors are shown in Fig. 2(top). Next, considering two non-zero attack sequences as $\tau^3_k \sim \mathcal{N}(0.2, 0.3)$ for $k \geq 60$ and $\tau^1_k \sim \mathcal{N}(0, 0.8)$ for $k \geq 40$, the distance measures v^i_k's over a sliding window of $T = 12$-step length are shown in Fig. 2(bottom). The figure clearly shows that the attacks affect the estimation error at all sensors. Setting two FARs $p_1 = 5\%$, $p_2 = 35\%$, the associated decision thresholds are $\theta_1 = 21$, $\theta_2 = 13.3$ via (16). From the figure, the less conservative threshold θ_2 reveals both attacks, while θ_1 only detects one and the other one remains stealthy most of the times.

5. CONCLUSION

We proposed an algorithm for simultaneous estimation of states and attack detection over a distributed sensor network. Using a windowed chi-square detector, every sensor is able to locally detect possible measurement anomalies causing the residuals to exceed an FAR-based threshold. As future research directions, the results in [5, 29] can be adopted to optimally locate the sensing nodes and design the network among the social sensors to reduce cost. Additionally, adopting the pruning strategies in [1, 2], one can change the social network structure and, in turn, tune its observability and information flow to improve estimation/detection properties.

6. REFERENCES

[1] M. Doostmohammadian, H. R. Rabiee, and U. A. Khan, "Centrality-based epidemic control in complex social networks," *Social Network Analysis and Mining*, vol. 10, pp. 1–11, 2020.

[2] P. Block, M. Hoffman, I. J. Raabe, J. B. Dowd, C. Rahal, R. Kashyap, and M. C. Mills, "Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world," *Nature Human Behaviour*, vol. 4, no. 6, pp. 588–596, 2020.

[3] H. Wai, A. Scaglione, and A. Leshem, "Active sensing of social networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 3, pp. 406–419, 2016.

[4] U. A. Khan and A. Jadbabaie, "Collaborative scalar-gain estimators for potentially unstable social dynamics with limited communication," *Automatica*, vol. 50, no. 7, pp. 1909–1914, 2014.

[5] S. Pequito, S. Kar, and A. P. Aguiar, "Minimum number of information gatherers to ensure full observability of a dynamic social network: A structural systems approach," in *IEEE Global Conference on Signal and Information Processing*. IEEE, 2014, pp. 750–753.

[6] M. Doostmohammadian, H. R. Rabiee, and U. A. Khan, "Cyber-social systems: modeling, inference, and optimal design," *IEEE Systems Journal*, vol. 14, no. 1, pp. 73–83, 2019.

[7] M. Doostmohammadian and U. Khan, "Graph-theoretic distributed inference in social networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 613–623, Aug. 2014.

[8] S. Kar and J. M. F. Moura, "Consensus + innovations distributed inference over networks: cooperation and sensing in networked systems," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 99–109, 2013.

[9] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3689–3704, 2018.

[10] P. Duan, Z. Duan, G. Chen, and L. Shi, "Distributed state estimation for uncertain linear systems: A regularized least-squares approach," *Automatica*, vol. 117, pp. 109007, 2020.

[11] X. He, X. Ren, H. Sandberg, and K. H. Johansson, "Secure distributed filtering for unstable dynamics under compromised observations," in *IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 5344–5349.

[12] M. Mansouri, M. N. Nounou, and H. N. Nounou, "Improved statistical fault detection technique and application to biological phenomena modeled by s-systems," *IEEE Transactions on Nanobioscience*, vol. 16, no. 6, pp. 504–512, 2017.

[13] H. Karimipour, A. Dehghantanha, R. M. Parizi, K. R. Choo, and H. Leung, "A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids," *IEEE Access*, vol. 7, pp. 80778–80788, 2019.

[14] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773–1786, 2015.

[15] B. Brumback and M. Srinath, "A chi-square test for fault-detection in Kalman filters," *IEEE Transactions on Automatic Control*, vol. 32, no. 6, pp. 552–554, 1987.

[16] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[17] J. Kim, C. Lee, H. Shim, Y. Eun, and J. H. Seo, "Detection of sensor attack and resilient state estimation for uniformly observable nonlinear systems having redundant sensors," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1162–1169, 2018.

[18] M. S. Chong, M. Wakaiki, and J. P. Hespanha, "Observability of linear systems under adversarial attacks," in *American Control Conference (ACC)*. IEEE, 2015, pp. 2439–2444.

[19] R. Tunga, C. Murguia, and J. Ruths, "Tuning windowed chi-squared detectors for sensor attacks," in *American Control Conference (ACC)*. IEEE, 2018, pp. 1752–1757.

[20] Z. Guo, D. Shi, K. H. Johansson, and L. Shi, "Optimal linear cyber-attack on remote state estimation," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 4–13, 2017.

[21] M. Doostmohammadian and N. Meskin, "Sensor fault detection and isolation via networked estimation: Full-rank dynamical systems," *IEEE Transactions on Control of Network Systems*, 2020.

[22] M. Deghat, V. Ugrinovskii, I. Shames, and C. Langbort, "Detection and mitigation of biasing attacks on distributed estimation networks," *Automatica*, vol. 99, pp. 369–381, 2019.

[23] B. Satchidanandan and P. R. Kumar, "Dynamic watermarking: Active defense of networked cyber–physical systems," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 219–240, 2016.

[24] M. Doostmohammadian and U. A. Khan, "Minimal sufficient conditions for structural observability/controllability of composite networks via kronecker product," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 78–87, 2019.

[25] M. Doostmohammadian and U. A. Khan, "On the characterization of distributed observability from first principles," in *2nd IEEE Global Conference on Signal and Information Processing*, 2014, pp. 914–917.

[26] U. A. Khan and A. Jadbabaie, "Coordinated networked estimation strategies using structured systems theory," in *49th IEEE Conference on Decision and Control*, 2011, pp. 2112–2117.

[27] P. E. Greenwood and M. S. Nikulin, *A guide to chi-squared testing*, John Wiley & Sons, 1996.

[28] M. Doostmohammadian, H. R. Rabiee, H. Zarrabi, and U. A. Khan, "Distributed estimation recovery under sensor failure," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1532–1536, 2017.

[29] M. Doostmohammadian, H. R. Rabiee, and U. A. Khan, "Structural cost-optimal design of sensor networks for distributed estimation," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 793–797, 2018.

Modified crop health monitoring and pesticide spraying system using NDVI and Semantic Segmentation: An AGROCOPTER based approach

Atharv Tendolkar[+] *Amit Choraria*[+] *Manohara Pai M M*[#] *Girisha S*[#] *Gavin Dsouza*[*] *K.S Adithya*[*]
Department of Electronics and Communication[+], Electronics and Instrumentation Engineering[*] &
Department of Information and Communication Technology[#]
Manipal Institute of Technology, MAHE Manipal

ABSTRACT

The technology in agriculture, can help farmers especially in the time of COVID pandemic, where there is shortage of labor and increasing demand for food. The technology solution can effectively and reliably improve crop yield through automated process and Agrocopter. The Agrocopter, an autonomous drone with modular systems and on-board image processing helps in holistic crop management throughout the farm. Agrocopter comes with targeted crop spraying, nutrient dropping and seed sowing modules, that can work in sync with the process of crop life cycle from sowing till harvesting. The drone with edge computing module performs periodic farm surveillance and plant health analysis using combination of NDVI (Normalized difference vegetation index) and semantic segmentation based classification to take targeted actions. It makes use of filter banks and SVM (Support Vector Machine) classifier algorithm to carry out pixel wise stitched image analysis to compute plant health indices in real time. Being very easy to operate and maintain, it can seamlessly be integrated into the farm systems and work along-side humans. It also has a completely modular design with plug and play architecture. What sets Agrocopter apart is its wide variety of applications, reliability and precision all at an affordable cost. Hence, Agrocopter is the perfect aerial farm assistant for today's farmer.

Index Terms— Autonomous, Drone, Edge Computing, Semantic segmentation, Support Vector Machines, NDVI

1. INTRODUCTION

The advancement in drone technologies and edge computing has driven innovative solutions for agriculture. The conventional agricultural practices rely on manual methods specially for monitoring crop health and pesticide spraying. These practices however became cumbersome during the COVID-19 pandemic times. This has motivated to develop Agrocopter after assessing the rising agricultural demands and hardships of farmers. It is designed to analyze the plant health and offer real time diagnosis to help improve yield and reduce the farmer's efforts. A systematic layered process is used to sense, analyze and act upon the field data [1]. The module basically encompasses a central flight control system along with an edge computing module on-board for image processing to perform targeted actions in real time. The analyzed data stored in the cloud system can be accessed by the farmer anytime, anywhere. Agrocopter is fully autonomous and can fly on a pre-planned path and perform instantaneous course correction around obstacles [2]. Being completely modular, the farmer has the liberty to attach any module and select the required mode which involves plantation, targeted crop spraying and solid nutrient dropping. All these functions will be carried out by the drone based on the time and schedule decided by the farmer. The on-board advanced camera system carries out smart on-site surveillance and performs advanced image processing including Semantic segmentation and NDVI (calculation of the amount of the visible and the reflected IR light from vegetation). The image processing algorithm uses in pixel wise reflectance analysis and SVM classifier to analyze crop health and accordingly decide to spray the pesticide. A smart wireless charging station with both portable and fixed configurations are provisioned for auto-docking and uploading mission data during the idle stage. The drone has failsafe features to prevent mishap and guarantee a better level of safety. Hence, this intelligent edge-based drone along with its IoT framework will go a long way in redefining agriculture.

2. LITERATURE REVIEW

At present, many aerial farm solutions rely on cloud-based satellite feed analysis. However, this offers low spatial resolution (30m), is expensive and requires clear skies [3]. Certain drone-based systems rely on wavelength capture using only NDVI parameter through modified cameras for farm analysis [4]. This however is less effective in areas of high biomass content and tends to amplify atmospheric and environmental noise. Alternative methods like Semantic segmentation of videos helps in scene understanding, thereby assisting in other automated video processing techniques like anomaly detection, object detection, event detection [5]. But, certain challenges like repeatability, computational complexity and labelling work offer challenges in this method. However, the Agrocopter uses a dual stage modified algorithm that utilizes NDVI and segmentation through an adaptive weightage computation based on environmental parameters captured through sensors. Current solutions offer complexity in operation, machinery, understanding and hence lead to higher cost [6]. The proposed solution

Agrocopter is a dovetail for the agricultural needs in terms of offering decision support and at the same time ensuring optimal cost. Modern day battery technology is not mature enough to give better flight time due to high power applications [7]. To ensure uninterrupted operations a smart charging dock is also discussed in the course of this paper. Hence the 'Agrocopter' provides a holistic, reliable and affordable aerial farm management system for the modern-day farmer.

3. WORKING AND MODELLING

3.1. System architecture

The basis of Agrocopter's design can be classified into a 4 level architecture which involves data capture, processing and transmission. Fig.1 shows its working with the underlining 4 layers.

1. Site : This is the portion of land under survey for the drone, in our case the crops on the farm. This is surveyed by onboard multispectral cameras.
2. Sensorics : The various sensors placed in the farm along with those on the drone form an integrated IoT network.
3. Drone : This layer involves the drone dynamics for generation of autonomous flight along with edge computing hardware.
4. Communication : In this layer the drone's navigation systems along with the data transmission systems to the cloud are involved.

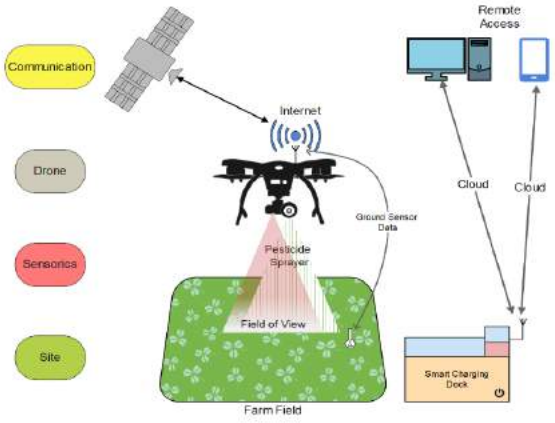


Figure 1. Layered Architecture

Meanwhile the layers are further divided to make optimal use of aerial imagery and modular plant management systems. Hence the Agrocopter is robust in design and working but at the same time easy to operate and seamlessly integrable.

3.2 Aerial System Development

The drone design was done keeping in mind its applications and user centricity. Some of the design parameters include flight time, altitude of flight, terrain and obstacle detection, payload capacity and seamless integration to perform edge computing and autonomous dynamics. The hexa-copter was selected keeping in mind the payload and degree of control. This provided a good balance between sophistication and value for money. The drone auto-stabilizes from LOS communication with the satellite and gets a 3-dimensional GPS lock for precision flight missions. The modular mechanisms were designed to be implemented using a simple plug and play architecture to reduce the learning curve for the farmer. Pesticide spray module[8] was designed to spray the liquid centrally under the drone to provide pinpoint precision. On the other hand, the Seed Box module [9] uses a flap controlled by the servo mechanism. The on-board edge computing unit processes all the operations including image analysis, actuation and decision support to activate plug and play unit such as pesticide sprayer, seed box module and camera vision system. Whereas autonomy is managed by the onboard Flight controller. The core of the Agrocopter is the Image processing unit with camera system this involves IR reflectivity scanning to give overall farm health status followed by semantic segmentation and classification for better precision. The generated masks are scaled, and the health ratio is obtained through the analyzer model as follows:

$$(\alpha * NDVI) + (\beta * Segmentation) = \eta \quad (1)$$

$$(\alpha * NDVI') + (\beta * Segmentation') = \eta' \quad (2)$$

Here, $\alpha + \beta = 1$ which suggests that the 2 methods will be set in a ratio based on environmental factors including weather, sunlight intensity and foliage configuration. In (1), we have the NDVI and scaled health ratio obtained from segmentation. In (2) the NDVI' and Segmentation' represent the ideal crop-specific health parameters obtained through repositories and scientific trial. This helps us to get final modified health ratios η and η'. The system now compares these two values and initiates trigger of sprayer in real time if ($\eta < \eta'$) which indicates unhealthy crop. This 2-stage algorithmic image analysis uses semantic segmentation further to give localized crop health status, thereby increasing the accuracy and optimal use of pesticides.

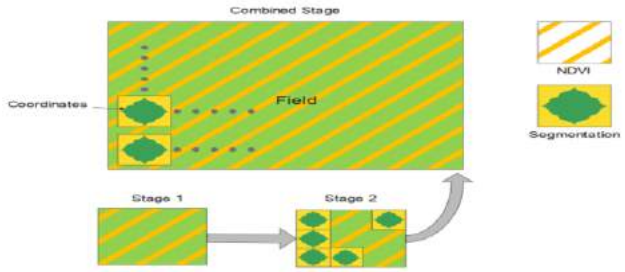


Figure 2. A modified 2 stage image analysis algorithm

Several flight modes can be assigned to the Agrocopter like land, loiter, stabilized, altitude-hold, auto, and return to Home. Auto missions can be pre-programmed and customized for the farmer. Hence, this gives an overview of how the Agrocopter (shown in Fig.2) was designed and the technology behind it.

3.3 Agro-Dock Development

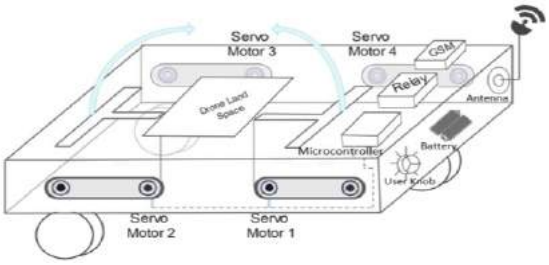
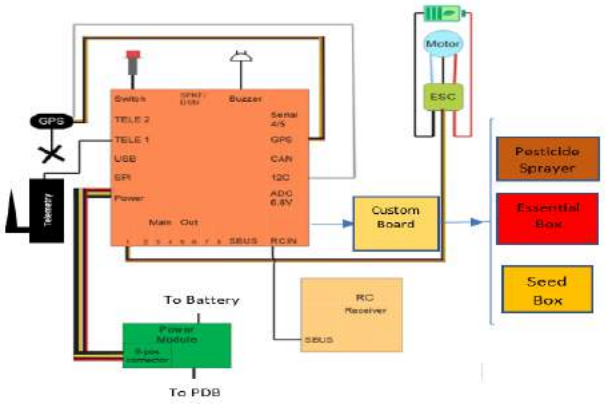


Figure 3. Smart Charging Station Schematic

The AgroCopter overcomes the challenges of modern battery technology for drones that affect the flight time. It comes with an in-house designed smart charging system named as Agrodock. This system is being offered in 2 variants, a fixed dock and a portable dock. The fixed dock is powered via standard power supply to safely charge the drone whereas the portable system uses automotive batteries, rechargeable every 10 days on an average thereby providing multiple recharge cycles for the drone. Apart from the wireless charging feature,[10] it also provides a connection to local Wi-Fi router or uses the mobile data through the GSM module encapsulated in the system. This collected data is relayed to cloud for storage and processing. The system consists of a two T-shaped energized coiled structure separated apart by the drone landing space. This landing platform consists of piezoelectric pressure sensors which makes the system aware of the absence or presence of the drone. When the drone is present on the land space, the dock would sense its presence which would trigger the smart relay that would raise the T-structure and make it move inwards using pulley mechanism. This inward movement would gently push the drone in position such that contact charging is enabled. The drone being powered by a Li-Po battery of 3S (8000mAh), gives a flight time of about 20-25 minutes. The microcontroller would trigger the relays to make the structure lay flat on the platform before the drone departs. Hence, the AgroDock is a perfectly safe and excellent complementary solution to wirelessly charge the Agrocopter.

3.4 System Integration

In the Flight Controller System, (Fig. 4) the series port of the 32-bit controller is connected to a GPS module for navigation using I2C protocol. The pesticide sprayer and seed box modules are connected to the custom board that drives these systems in accordance with the plant health algorithm and flight control system. For our application, the voltage is kept at a maximum of 12.6V generating power which is equally distributed to 6 electronic speed controllers that drive 10-inch propellers via 920kV high torque brushless DC motors. A hardware safety switch and buzzer are also connected to the controller to prevent any accidental arming of the drone. Agrcopter generates 7 kg of thrust and can lift payloads up to 3 kg while providing a decent flight ratio giving it better flight dynamics and agility.



4. SYSTEM IMPLMENTATION

The Agrocopter makes use of several sensors, multi-spectral cameras and custom-built modules that work in sync to ensure targeted operation and optimized system performance. The following sub-systems describe the individual operation, working and contribution in Agrocopter.

4.1 Agro Systems

The Agrocopter has been designed to be completely modular with plug and play systems for agriculture purposes. As shown in Fig. 5 the major systems on board the drone can be discussed as follows :

1. Pesticide Sprayer : This in-house built sprayer provides targeted spraying on crops to mitigate diseases. This is driven by a 5V mini pump with 80 L/H capacity triggered by the edge-based system, in response to poor plant health as analyzed by image processing.
2. Payload box: This can be used to carry up to 500 grams extra required materials, sensors or first aid throughout the farm [11].
3. Seed Box : During plantation season, the drone can be used for systematic positional periodic seed dispersing driven by timed waypoint-based servo release mechanism. This module can be used for solid pesticide as well.

Drone testing was carried out in a field in Manipal India with coordinates (13.344529651512422, 74.79392134764375)

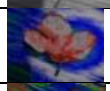
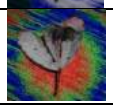
Figure 5. Agro Modules and Camera in Action

4.2 Camera System

The camera used in the Agrocopter is a 1200 TVL CMOS FPV camera with 2.8mm lens that was modified to allow IR imaging. The multi-spectral image data captured is used during NDVI and semantic segmentation computation [12]. This live aerial feed as shown in Fig. 5 is captured and a composite farm image is stitched on Mission Planner software from where it is sent to the cloud, and analyzed using 2 different methods as follows:

NDVI : In the first method, the image is analyzed pixel wise and each pixel for the crop images is assigned a value between 0 to 1 according to the NDVI scale which generates result as shown in Table 1. NDVI is measure of the difference in reflectance of visible and near-infrared light from the vegetation, expressed as a ratio. This helps to analyze rudimentary plant health and take actions accordingly.

Table 1. NDVI results

Color Image	NDVI image	NDVI value	Health
		0.80	Healthy
		0.32	Moderately healthy
		0.15	Unhealthy

Semantic segmentation : The second parallel process involves assigning a class to every individual pixel in an image and classify it into healthy and unhealthy classes based on a machine learning algorithm. It uses an architecture consisting of a feature descriptor and classifier. In general, there exists color and texture variations between healthy and unhealthy crops as shown in Table 2. Texture and color features are considered as a feature descriptor. Texture features are calculated using filter banks which consist of multiple filters that represents different patterns. The filter bank considered in the present study has 17 filters of Gaussian and Laplacian. The Gaussian filter is applied at varying scales by setting standard deviation (sigma) to 1, 2 and 4. These filters are applied to R, G and B color channels. X and Y derivative of Gaussian filter is applied with sigma value set to 2 and 4. Laplacian filter is applied with sigma value set to 1, 2, 4 and 8, whiles RGB and LAB color features are considered along with the texture features Support Vector Machine (SVM) classifier is trained to map the feature vector to two classes. The parameters of SVM classifier are identified by utilizing 10-fold cross validation. In the present study, the gamma value of SVM is set to 0.0001 and the C parameter which provides penalty is set to 10. For every pixel, texture and color features are extracted and classified by utilizing SVM classifier. Based on Texture and pixel wise probability of green color intensity, we compute the individual pixel health status. This is then converted to a mask based on if the values are greater or smaller in comparison to the healthy threshold. The observed masks as in Table 2, help to get a net pixel area ratio (white area / total image area), which forms the health ratio of the complete image. This is then scaled to fit the weightage-based parameter model equation.

Table 2. Sample Field Images with segmentation masks

Color Image	Segmented Mask	Health Ratio	Scaled Health Ratio	Health
		0.0729	0.729	Healthy
		0.0646	0.646	Moderately Healthy
		0.0256	0.256	Unhealthy

The performance of the semantic segmentation model is evaluated by calculating standard metrics, i.e., precision, recall and f1-score. The algorithm gives a precision of 0.85 and a recall of 0.8 on the training dataset. Further, the performance was evaluated against samples consisting of only one class (healthy or unhealthy) as shown in Table 3.

Table 3. Classification Report

Metric	Training set	Healthy	Unhealthy
Precision	0.85	1.0	1.0
Recall	0.81	0.91	0.86
F1-score	0.79	0.95	0.92

Although NDVI is computed without environmental parameters, it is non-linear and is sensitive to crop background brightness. Whereas Semantic segmentation can overcome the above issues but has computational complexity and poor scaling. The generated models and masks are scaled by factors (α and β) that give the modified and accurate result for crop health analysis. Hence, this dual step-based approach not only gives a more robust solution, but also provides a reliable method of monitoring the health of the crops with increased accuracy.

5. CONCLUSION

The above sections give an idea of the blueprint of the Agrocopter's design along with its associated image analytics technology capabilities. The implementation of modified crop health analysis based on sub-weightage allocation through environmental factors helps to strategically exploit the benefits of segmentation and NDVI methods and prevents their drawbacks. The modular approach along with pesticide sprayer, seed module and payload box offer holistic service and aid agriculturists. Moreover, it is seen that the drone is able to remotely manage farms and mitigate crop health issues real time thereby maximizing profits and reducing efforts for farmers across the globe. Agrocopter will surely go a long way in providing the modern farmer a low-cost, reliable and intelligent aerial farm assistant system.

5.1. Future Development

The Agrocopter is soon to be industry ready. It is fully capable of carrying out autonomous missions and farm image surveillance with crop management. The on-board controllers and systems are capable of integrated NDVI imaging and segmentation on the edge. The subsequent development will involve turning the Agro-Dock idea into a reality to ensure seamless charging and data porting. Work is also going on to make it water and dust proof to survive the harsh scenarios in farms anywhere across the globe. It can be used to analyze ripe fruits and pick them during the season thereby reducing labor costs. The image analytics can understand the diseases faced by the plant and spray the exact required chemical from its on board payload catalog. Moreover, being completely modular, it can be used not just in the agriculture space but also for geospatial mapping, search and rescue, medical emergencies, food and goods deliveries and much more. Having a robust design and payload lifting ability, it can also carry heavy cameras, parcels and equipment. Hence, it is seen that the applications of Agrocopter are endless, and it will surely go a long way in the holistic and sustainable progress of mankind.

REFERENCES

[1] Saha AK, Saha J, Ray R, Sircar S, Dutta S, Chattopadhyay SP, Saha HN. IOT-based drone for improvement of crop quality in agricultural field. In2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC) 2018 Jan 8 (pp. 612-615). IEEE.

[2] Koh LP, Wich SA. Dawn of drone ecology: low-cost autonomous aerial vehicles for conservation. Tropical conservation science. 2012 Jun;5(2):121-32.

[3] Kulbacki M, Segen J, Knieć W, Klempous R, Kluwak K, Nikodem J, Kulbacka J, Serester A. Survey of drones for agriculture automation from planting to harvest. In2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES) 2018 Jun 21 (pp. 000353-000358). IEEE.

[4] Mahajan U, Raj B. Drones for normalized difference vegetation index (NDVI), to estimate crop health for precision agriculture: A cheaper alternative for spatial satellite sensors. In Proceedings of the International Conference on Innovative Research in Agriculture, Food Science, Forestry, Horticulture, Aquaculture, Animal Sciences, Biodiversity, Ecological Sciences and Climate Change (AFHABEC-2016), Delhi, India 2016 Oct 22 (Vol. 22).

[5] S. Girisha, U. Verma, M. M. Manohara Pai and R. M. Pai, "UVid-Net: Enhanced Semantic Segmentation of UAV Aerial Videos by Embedding Temporal Information," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4115-4127, 2021, doi: 10.1109/JSTARS.2021.3069909.

[6] Mogili UR, Deepak BB. Review on application of drone systems in precision agriculture. Procedia computer science. 2018 Jan 1;133:502-9.

[7] Raciti A, Rizzo SA, Susinni G. Drone charging stations over the buildings based on a wireless power transfer system. In2018 IEEE/IAS 54th Industrial and Commercial Power Systems Technical Conference (I&CPS) 2018 May 7 (pp. 1-6). IEEE.

[8] Wen S, Han J, Ning Z, Lan Y, Yin X, Zhang J, Ge Y. Numerical analysis and validation of spray distributions disturbed by quad-rotor drone wake at different flight speeds. Computers and Electronics in Agriculture. 2019 Nov 1;166:105036.

[9] Fortes EP. Seed plant drone for reforestation. The Graduate Review. 2017;2(1):13-26.

[10] Choi CH, Jang HJ, Lim SG, Lim HC, Cho SH, Gaponov I. Automatic wireless drone charging station creating essential environment for continuous drone operation. In2016 International Conference on Control, Automation and Information Sciences (ICCAIS) 2016 Oct 27 (pp. 132-136). IEEE.

[11] Riananda DP, Nugraha G, Putra HM, Baidhowi ML, Syah RA. Smart pulley workflow in delivery drone for goods transportation. InAIP Conference Proceedings 2020 Apr 21 (Vol. 2226, No. 1, p. 060010). AIP Publishing LLC.

[12] Yanowitz SD, Bruckstein AM. A new method for image segmentation. Computer Vision, Graphics, and Image Processing. 1989 Apr 1;46(1):82-95.

LOCAL, GLOBAL AND SCALE-DEPENDENT NODE ROLES

Michael Scholkemper, Michael T. Schaub

Department of Computer Science, RWTH Aachen University, Germany

ABSTRACT

This paper re-examines the concept of node equivalences like *structural equivalence* or *automorphic equivalence*, which have originally emerged in social network analysis to characterize the role an actor plays within a social system, but have since then been of independent interest for graph-based learning tasks. Traditionally, such exact node equivalences have been defined either in terms of the one-hop neighborhood of a node, or in terms of the global graph structure. Here we formalize exact node roles with a scale-parameter, describing up to what distance the ego network of a node should be considered when assigning node roles — motivated by the idea that there can be local roles of a node that should not be determined by nodes arbitrarily far away in the network. We present numerical experiments that show how already "shallow" roles of depth 3 or 4 carry sufficient information to perform node classification tasks with high accuracy. These findings corroborate the success of recent graph-learning approaches that compute approximate node roles in terms of embeddings, by nonlinearly aggregating node features in an (un)supervised manner over relatively small neighborhood sizes. Indeed, based on our ideas we can construct a shallow classifier achieving on par results with recent graph neural network architectures.

***Index Terms*—** role extraction, node roles, graph learning, graph neural networks

1. INTRODUCTION

Networks have become a powerful abstraction to understand a range of complex systems [1, 2]. To comprehend such networks we often seek patterns in their connections, e.g., densely-knit clusters or core-periphery structure, which would enable a simpler or faster analysis of such systems. The concept of node roles or node equivalences, originating in social network analysis [3] is another of those patterns used to simplify complex networks. The questions underpinning the detection of node roles are i) what nodes serve what function within the network? and ii) which nodes are similar in functionality?

The intricacy of node role extraction derives from the lack of a clear definition of the role of a node in mathematical terms. Traditional approaches, put forward in the context of social network analysis [4] consider *exact node equivalences*, based on structural symmetries within the graph structure. The earliest notion is that of *structural equivalence* [5], which assigns the same role to two nodes if they are adjacent to the exact same nodes. Another definition is that of *automorphic equivalence* [6], which states that nodes are equivalent if they belong to the same automorphism orbits. Closely related is the idea of *regular equivalent* nodes [7], defined recursively as nodes that are adjacent to equivalent nodes.

We acknowledge partial funding from Ministry of Culture and Science of North Rhine-Westphalia (NRW Rückkehrprogramm) and the Excellence Strategy of the Federal Government and the Länder.

Interestingly, essentially all of these approaches determine the role of the node either purely on the direct neighborhood of a node, or in terms of the "global neighborhood", i.e., the relative position of a node within the whole graph. Motivated by this observation, here we re-consider such exact node equivalences and introduce a scale parameter for the definition of exact node equivalences, and discuss algorithms that are able to find the thus defined (local) node roles. The overall idea here is that rather than having a single, fixed scale for defining a node role, there may be different relevant scales for node roles, depending on the problem at hand — indeed one may argue that, e.g., in a social network context the role of a node should not be dependent on nodes potentially arbitrary far away.

Related literature Apart from the above mentioned ideas of *exact* node equivalences, there exists a large number of works on role extraction, which focus on finding nodes with *similar* roles (though not identical). Many of these methods are based on computing feature vectors or embeddings of nodes. Based on these embedding we can then calculate pair-wise similarities between nodes and cluster them into groups. The overview article [8] puts forward three categories: First, graphlet-based approaches [9, 10, 11] use the number of graph homomorphisms of small structures to create node embeddings. This retrieves extensive, local information such as the number of triangles a node is part of. Second, walk-based approaches [12, 13] embed nodes based on certain statistics of random walks starting at each node. Finally, matrix-factorization-based approaches [14, 15] find a rank-r approximation of a node feature matrix ($F \approx MG$). Then, the left side multiplicand $M \in \mathbb{R}^{|V|\times r}$ of this factorization is used as a soft assignment of the nodes to r clusters.

Jin et al. [16] provide a comparison of many of such node embedding techniques in terms of their ability to capture exact node roles such as structural, automorphic and regular node equivalence. Detailed overviews of (exact) role extraction and its links to related topics such as block-modelling are also given in [17, 18]. The idea of node roles is also similar to community detection [19]. However, whereas the locality of connections is often a central factor for community detection, this is less of a consideration for node roles. Indeed, nodes that are far apart or are even part of different connected components, can have the same role [8].

Contributions and outline. We provide a fresh perspective on the definition of exact node roles which has certain parallels in recent developments in graph-based machine learning techniques such as graph neural networks. First, we formalize exact node roles with a scale parameter, describing up to what distance the graph structure surrounding a node should be considered when assigning node roles. Second, we re-frame the well-known Weisfeiler Lehman (WL) algorithm, and show how it can be interpreted in terms of a relaxation of computing such local node roles. We then provide a different algorithm that solves the scale-dependent node embedding, using an alternative problem relaxation, which may be interpreted as a dual of the WL algorithm. Finally, we show that *shallow* roles provide enough information to achieve high performance on node classifica-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

tion tasks, and compare these results based on local node roles with the performance of recent graph neural network architectures.

2. NOTATION AND BACKGROUND

Graphs. An undirected *graph* G consists of a vertex-set V and edge-set $E \subseteq \{\{u, v\} | u, v \in V\}$ describing relations between the vertices. Given a subset $V' \subseteq V$, the *graph induced by* V' is defined as the graph with vertex set V' and edge set $E' = \{\{u, v\} \in E | u, v \in V'\}$. For a vertex v, we define its *neighbourhood* as the set $N(v) = \{x | \{v, x\} \in E\}$. The *k-hop-neighbourhood* $N^k(v)$ is the set of nodes reachable from v in at most k steps.

Colorings and refinements. We define a graph *coloring* as a function $c : V \to \{1, \dots, \mathcal{C}\}$ which maps each node to one out of $\mathcal{C} \in \mathbb{N}$ many colors. The classes of a coloring are the node sets associated to the same color $C_i = \{v \in V \mid c(v) = i\}$. The partition associated to these classes induces an equivalence relation $\sim_c$ among the vertices. We say a coloring c *refines* another coloring c', written $c \sqsubseteq c'$, if for all node $u, v \in V$ a different color assignment $c'(u) \neq c'(v)$ under coloring c' implies a different assignment $c(u) \neq c(v)$ under coloring c. Hence, the partition induced by the color-classes of c is a subpartition of the partition induced by c'. Two colorings c, c' are *equivalent*, written $c \equiv c'$, if they refine each other ($c \sqsubseteq c'$ and $c' \sqsubseteq c$). This implies that $\sim_c$ and $\sim_{c'}$ define the same relation. A *refinement* is an iterative algorithm, that produces at each iteration t a coloring c^t that refines the previous coloring, i.e., $c^{t+1} \sqsubseteq c^t$.

Isomorphism and orbits. An isomorphism between two graphs G, G' is function that maps all vertices from one graph to the other graph, while preserving adjacency relationships and coloring. Formally, given colored graphs $(G, c), (G', c')$, an isomorphism is a bijection $\pi : V(G) \to V(G')$ such that (i) $\{u, v\} \in E(G) \Leftrightarrow \{\pi(u), \pi(v)\} \in E(G')$, and (ii) $c(v) = c(\pi(v))$ for all vertices. If such an isomorphism exists between graphs G and G', we say that these graphs are *isomorphic* $G \cong G'$. An *automorphism* is an isomorphism from a graph to itself $\pi' : G \to G$. The *orbit* $\text{orb}_G(v)$ of a vertex v is the set of all vertices u for which there exists an automorphism π' such that $\pi'(v) = u$.

Unravellings. Let (G, c) be a colored graph and $v \in V$. The *node-unidentified unravelling* $\mathcal{U}^d(v)$ of depth d rooted at node v is the tree defined as follows. The vertex set of $\mathcal{U}^d(v)$ is the set of walks of length at most d starting at v. Furthermore two nodes $w_1 = (v, x_1, ..., x_n)$ and $w_2 = (v, y_1, ..., y_{n-1})$ are connected if $x_i = y_i$ for $1 \leq i \leq n - 1$, i.e., if the walk corresponding to w_1 is simply an extension by one node of the walk corresponding to w_2. This definition induces a natural equivalence relation between two nodes $u \sim^d_{\mathcal{U}} v$, where nodes are equivalent if and only if $\mathcal{U}^d(u) \cong \mathcal{U}^d(v)$.

The *node-identified unravelling* $\mathcal{I}^d(v)$ of depth d rooted at node v is defined analogously as above, with one addition: Each node $(v, x_1, ..., x_n)$ is identified by the vertex x_n at the end of the walk, i.e. $\text{id}((v, x_1, ..., x_n)) = x_n$. For two nodes u, v to be equivalent $u \sim_{\mathcal{I}} v$, there must exist an isomorphism σ certifying $\mathcal{U}^d(u) \cong \mathcal{U}^d(v)$ such that $\pi : \text{id}(x) \to \text{id}(\sigma(x))$ is a well-defined bijection. In other words, all nodes in $\mathcal{U}^d(u)$ with the same id i are consistently mapped onto nodes with the same id j in $\mathcal{U}^d(v)$ by σ (but it is not required that $i = j$). The motivation underpinning this definition is that a consistent isomorphism σ between the unravellings induces a local isomorphism π between the graphs (see Prop. 1, Fig. 1).

3. LOCAL, GLOBAL AND SCALE DEPENDENT ROLES

Though assigning roles to a node in a network is an intuitively simple idea, there is an inherent tension in the definitions of roles: on the one hand we want to identify roles that help us to comprehend the system on a global level, on the other hand the semantics of the data encoded in the network is often much more local.

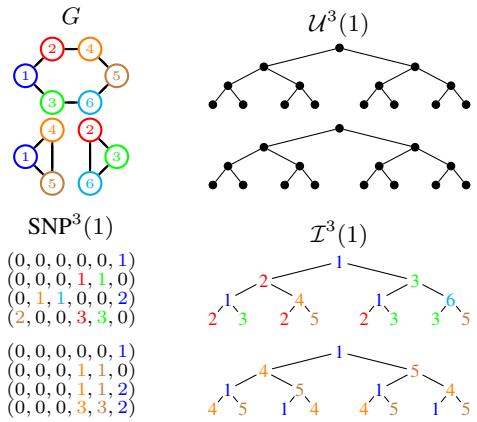


Fig. 1: **Illustration of different d-role definitions.** The figure shows two graphs, the identified and unidentified unravellings as well as the SNP-embeddings of node 1 in the respective graph. The 0- and 1-roles of the nodes in the graphs are the same, but for $d \geq 2$, the d-roles are different. There are 5 distinct ids in the upper unravelling, but only 3 distinct ids in the lower one, thus there exists no bijection that is consistent with respect to the ids. This is reflected in the SNP-embedding whose roles are also different for $d \geq 2$. The d-wl-roles, however, are the same for all nodes and all d. (Colors are only used for visual clarity, they are interchangeable.)

To illustrate this, consider a social network. Intuitively, the direct neighbourhood of a node defines its sphere of influence and thus should be considered when analysing which role a node plays. Applying this argument recursively, the same is true for the second hop neighbourhood of a node. Following this train of thought, we may argue that automorphism orbits define nodes roles in the most natural way: these are the finest possible role definitions which are isomorphism preserving and thus independent of the ordering of the node labels. Yet, with every hop, the relative influence of a node within a social network is bound to decrease. Automorphic equivalence may thus lead to undesirable node assignments, as any minor asymmetry in the graph far away from a node could influence its role. We may thus prefer a more local definition of node roles.

However, a global view on node roles may indeed be desired in other scenarios. For instance, in molecular biology we may encode a protein as a graph of amino acids and their interactions. However, the overall protein structure can be shaped by even far apart amino acids, local changes in the amino acids can influence the functionality of reaction sites and the whole protein.

An inherent scale is also present in most approaches to extract approximate node roles [8, 18]. For computational reasons, most methods calculate feature vectors of local node statistics, which are then aggregated via clustering. For example, graphlet density calculation [9, 10], or random walk based statistics [12, 13] result in highly local features for every node, which are then used to compute approximate node roles. However, the scale of the extracted features is thus fixed and not adaptable to the problem at hand. In fact in some cases, there may be multiple sensible node-role assignments, depending on the scale of the problem one is interested in.

We therefore propose a formal definition of node roles that combines these demands through a scale-parameter.

Definition 1. *The d-depth node role (d-role) is given by the structural equivalence of the node-identified unravellings. Equivalently, two nodes u, v have the same d-roles if and only if $u \sim_{\mathcal{I}}^{d} v$.*

Thus, the d-role of a node is local for small d, since only nodes in the d-hop neighbourhood are used to obtain the roles. In contrast, the d-role can also include more global features, if d is chosen sufficiently large. This mirrors developments in community detection [19], the analysis of mixing patterns [20], or in centrality measures such as Katz centrality [21], the scale of interest can often be adjusted via a parameter.

Our first shows that for large d, our node role definition coincides with automorphic equivalence.

Proposition 1. *Let G_1, G_2 be two (colored) graphs of the same size and $u \in V(G_1), v \in V(G_2)$. For $d = \max(|V(G_1)|, |V(G_2)|)$, it holds that $u \sim_{\mathcal{I}}^{d} v$ if and only if $u \in orb_{G_1 \cup G_2}(v)$.*

Proof. For the backward direction, suppose $u \in \text{orb}(v)$ and let π be the automorphism certifying this. Then the permutation

$$\sigma((v, x_1, ..., x_n)) \coloneqq (\pi(v), \pi(x_1), ..., \pi(x_n))$$

proves that $\mathcal{U}^d(u) \cong \mathcal{U}^d(v)$ and $\text{id}(x) \to \text{id}(\sigma(x)) = \pi(\text{id}(x))$ is well-defined and bijective.

For the other direction, suppose there exists a σ that proves $\mathcal{U}^d(v) \cong \mathcal{U}^d(u)$, such that $\pi : \text{id}(x) \to \text{id}(\sigma(x))$ is well-defined and bijective. Let $\tilde{E}(v) = E(G_1[N^d(v)])$, i.e. the set of all edges reachable from v. Consider the edge $\{x_1, x_2\} \in \tilde{E}(v)$. Then there exist both a walk w_1 starting at v and ending in x_1 as well as a walk w_2 that is the same as w_1 but extended by x_2. w_1 and w_2 are neighbours in $\mathcal{U}^d(v)$. Thus, $\sigma(w_1)$ and $\sigma(w_2)$ must be neighbours in $\mathcal{U}^d(u)$, and so $\pi(x_1) = \text{id}(\sigma(w_1))$ and $\pi(x_2) = \text{id}(\sigma(w_2))$ are neighbours in G_2.

By symmetry of the argument, taking $\{x_1, x_2\} \in \tilde{E}(u)$ and using the inverses of π and σ, yields that π is an isomorphism between $\text{comp}(v)$ and $\text{comp}(u)$. Thus extending π by the identity for all nodes in $V(G_1 \cup G_2) \backslash (N^d(v) \cup N^d(u))$ yields an automorphism on $G_1 \cup G_2$. □

Hence, our definition of node roles is rigorous in the sense that for large d it is as expressive as automorphic equivalence, as is well defined locally in the sense that $u \sim_{\mathcal{I}}^{d} v$ induces a local isomorphism between $G[N^{d-1}(v)]$ and $G[N^{d-1}(u)]$. However, the above criterion is hard to check computationally. Specifically, a direct consequence of the above proposition is that computing the d-roles is at least as hard as solving the graph isomorphism problem. The fastest known algorithm for this requires quasi-polynomial time [22], which is intractable for most large problems. For fixed d, the problem may be computable in polynomial time, but it is still linked to local isomorphism. In the following, we therefore propose relaxations of our local role definition that allow efficient computation.

4. RELAXATIONS

In this section we provide two relaxations of the problem of computing d-roles that can be computed efficiently. These relaxations are dual in the following sense. In the first case, we drop the node identifiers in the (identified) unravellings and use node-unidentified unravellings instead. In the second case, we neglect the detailed knowledge about the (identified) unravelling structure, but keep the information of the node identities at every step. We start with the first case, which is closely connected to some well known graph-isomorphism algorithms.

Definition 2. *The d-depth WL node role (d-wl-role) is given by the equivalence induced by the node-unidentified unravellings.*

The equivalence relation $\sim_{\mathcal{U}}^{d}$ given by this relaxation coincides with the coloring after d iterations of the so-called color refinement algorithm, which is also known as the 1-dimensional Weisfeiler Lehman algorithm [23]. Starting from a constant initial coloring, this algorithm iteratively computes the node colors according to the following formula:

$$c^{t+1}(v) = \text{hash}\left(c^t(v), \{\{c^t(x) | x \in N(v)\}\}\right)$$

where hash is an injective hash-function, and $\{\{\cdot\}\}$ denotes a multiset (a set in which elements can appear more than once). With every iteration the algorithm aggregates information from its neighbours who, in turn, have aggregated information from their neighbours previously and so on. After d iterations, the colors of nodes have information about nodes that are at most distance d apart. This information is exactly captured in the node-unidentified unravelling:

Proposition 2. *Let G be a graph and let c^d be the colors of the color refinement algorithm after d iterations. Then the equivalence relation induced by the coloring c^d corresponds to the equivalence relation induced by the d-step unravelling $\sim_{c^d} \equiv \sim_{\mathcal{U}}^{d}$.*

Proof sketch. Consider the color assignment $c^1(v)$ in the first step for a node v. It encodes the degree of v as the injective hash-function enables the reconstruction of the multi-set of neighboring colors. Since all colors are initially the same, this multi-set has one element with muliplicity of $\deg(v)$. Following the same logic, in the second iteration the color $c^2(v)$ encodes how many neighbors have which degree. Iterating this idea results in the above claim. □

While not phrased in terms of node roles, the result of Prop. 2 is essentially known in the literature, which is why we only sketch the proof here. Indeed, color refinement is a well-known algorithm. For example, [24] showed that it distinguishes almost all random graphs and there are also results linking the expressive power to fragments of logic [25] and, more recently, graph neural networks [26, 27]. Nonetheless, thinking in terms of d-roles provides a fresh look on both the algorithm and the problem at hand. Typically, we care only about the (final) *stable* coloring given by the algorithm, which yields the coarsest equitable partition of the graph. However, here we are interested in the preliminary coloring of depth d as it provides us with (relaxed) node roles. Moreover, the computation of these d-wl-roles is quite efficient: k iterations of color refinement can be computed in time $\mathcal{O}(k \cdot |V(G)| \cdot \deg(G) \cdot \log(\deg(G)))$, where $\deg(G)$ is the maximum degree of G. While this efficiency is our main motivation to use this algorithm, as our experiments show in the following section, the resulting d-wl-roles are very effective for graph analysis tasks as well.

Let us now consider a second, dual relaxation of d-roles, where instead of removing the node identifiers, we instead remove the structure in the unravelling. More specifically, we count the multiplicity of unique identifiers at every depth of the unravelling and use this data to define the node role. The exact information which node is connected to which other node is thus generally neglected in our node assignment. Still, the expressivity of the assigned node roles is empirically comparable to color refinement. Moreover, this second procedure performs much better on regular graphs, where color refinement is known to fail, and provides an intuitive and simply explainable representation.

Definition 3. *The d-depth SNP node role (d-snp-role) is given by the multi-set of identifiers in the unravelling at every level up to d.*

From the definition of the node-identified unravelling, we see that counting the identifiers at level d is equivalent to counting the number of walks of length d from the root to all nodes reachable in exactly d steps. It is thus closely tied to the v-th row of the adjacency matrix power $(A^d)_{v,\text{-}}$. However, this row-vector will depend on the ordering of the nodes. Instead, we consider the node embedding:

$$\text{SNP}^d(v) = \text{lex-sort}\left(\begin{bmatrix} A^0_{v,\text{-}} \\ \vdots \\ A^d_{v,\text{-}} \end{bmatrix}\right)$$

where lex-sort sorts matrix columns lexicographically. We call this *sorted Neighbourhood Propagation* (SNP) embedding. The SNP embedding is isomorphism-invariant, since a permutation of the adjacency matrix preserves the multi-set of columns of the matrix input to the lex-sort and lex-sort, in turn, uniquely determines the location of any column up to equality. Computing the embedding takes time $\mathcal{O}(k \cdot |V(G)|^{2.37})$, in general though sparse matrix multiplication speeds up the process for typical graphs. We thus have a second relaxation of node-roles based on SNP, where we say two nodes have the same role if their SNP embedding is the same.

5. APPLICATIONS/EXPERIMENTS

The following section presents the results of two computational experiments. First, we examine the local roles obtained from the WL and SNP algorithm, and see how they can be employed for the analysis of small (188 graphs), medium (600 graphs) and large (2000 graphs) data-sets, namely MUTAG [28], ENZYMES [29] and NCI1 [30]. In all three datasets, nodes are annotated with attributes and there exists a target label for each graph. A detailed description of the datasets can be found in [31]. In our second experiment, we compare the results of training a multi-layer perceptron (MLP) classifier using the SNP-embedding as inputs with directly training a graph neural network. Throughout both experiments we use node-level and graph-level targets. Code will be made available here.

Experiment 1. We investigate the utility of d-wl-roles and d-snp-roles for a node classification task. To this end, Figure 2 shows the number of distinct color classes/embeddings that both algorithms find at each depth d, divided by the number of nodes in the datasets. As can be seen, the local roles are highly discriminitative, already at small depths. To determine whether these local node roles are useful, we check whether the node roles classes are correlated with some of the node attributes and find that this is indeed the case. Figure 2 shows the overlap score obtained, if we simply matched the nodes within a role-class with the most frequently occurring label. The overlap score here is defined as $(a-b)/(1-b)$, where a is the accuracy of the assignment of node roles to the node label, and b is a baseline accuracy. We use the accuracy at depth 0 as the baseline, i.e., the accuracy of simply always guessing the most probable label.

Remarkably, a depth of 3–4 suffices to obtain an overlap score of more than 90% for the three data-sets except NCI1. The latter requires a neighborhood depth of 7 or more to achieve a similar score. Obviously, both of our role-detection methods are not directly suitable as node label classifiers, as they do not generalize to unseen data. Rather, our results show that the information needed to obtain high accuracy scores in the here considered node classification tasks is present in the near surroundings of the nodes — without specifying how to exploit it.

Experiment 2. In the second experiment, we address this aspect by comparing the GIN graph neural network architecture [27] — provably one of the most powerful GNNs, while still permutation invariant — with a multi-layer-perceptron (MLP) classifier that is given the SNP embedding as input. The GIN uses a 2-layer MLP to update the embeddings and a 3-layer MLP as the final classification layer. For comparability, the final classifier in the SNP approach has the same size. Table 1 shows the mean accuracy on the test sets for 10 separate 10-fold cross-validations for each model. In the hyperparameter search, the SNP classifier was only allowed a depth of $3-4$, whereas the GIN was allowed a depth of up to 10.

All in all, the SNP classifier is competitive with the GIN on the node-level and on the graph-level. The comparison with the uninformed GIN- shows that the SNP classifier has the edge when it comes to extracting information — even if the GIN was significantly deeper. Interestingly, the SNP classifier tended to overfit — reaching $> 99\%$ accuracy on nearly all training datasets — whereas the GIN tended to achieve similar accuracy on the train-splits as on the test-splits.

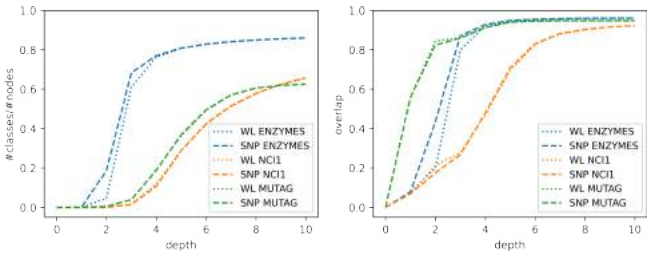


Fig. 2: **Disriminative power and utility of local roles.** Number of roles relative to the dataset size (left) and overlap (right) versus depth d. Dotted lines indicate the d-wl-role, dashed lines the d-snp-role.

		MUTAG	ENZYMES	NCI1
Node-Level	SNP-	$\mathbf{96.7 \pm 0.1}$	$\mathbf{65.8 \pm 0.1}$	$\mathbf{86.4 \pm 0.0}$
	GIN-	96.1 ± 0.1	58.4 ± 0.2	$\mathbf{86.7 \pm 0.1}$
Graph-Level	SNP-	$\mathbf{96.5 \pm 0.5}$	51.0 ± 0.6	79.3 ± 0.1
	GIN-	94.3 ± 0.3	31.8 ± 0.7	75.6 ± 0.2
	GIN+	94.3 ± 0.5	$\mathbf{60.3 \pm 0.7}$	$\mathbf{82.0 \pm 0.3}$

Table 1: **Accuracy (in $\%$) of the SNP classifier and the GIN on the node-level and the graph-level tasks.** $+$ and $-$ indicate whether the model had access to the node attributes or not. We report the mean over 10 crossvalidation runs along with the standard deviation.

6. CONCLUSION

We formalized the idea of scale-dependent node roles, presented two algorithms to compute such roles and demonstrated their representational power for certain graph learning tasks. Many recent developments in machine-learning see a strive toward deep classifiers. However, certain architectures, such as the GCN architecture [32], are not suited to be very deep [33, 34]. Our experiments indicate that some graph learning tasks may indeed not require deep features (at least in terms of role depth) and that simple classifiers based on local node roles can yield surprisingly competitive results. Possible future directions include work on efficient algorithms to compute d-roles, incorporating external node features into the SNP-embedding, or establishing other similarity scores based on these ideas.

7. REFERENCES

[1] Mark Newman, *Networks*, Oxford University Press, 2018.

[2] Steven H Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[3] David Knoke and Song Yang, *Social network analysis*, Sage Publications, 2019.

[4] Ulrik Brandes, *Network analysis: methodological foundations*, vol. 3418, Springer Science & Business Media, 2005.

[5] Francois Lorrain and Harrison C White, "Structural equivalence of individuals in social networks," *The Journal of mathematical sociology*, vol. 1, no. 1, pp. 49–80, 1971.

[6] Martin G Everett and Stephen P Borgatti, "Regular equivalence: General theory," *Journal of mathematical sociology*, vol. 19, no. 1, pp. 29–52, 1994.

[7] Douglas R White and Karl P Reitz, "Graph and semigroup homomorphisms on networks of relations," *Social Networks*, vol. 5, no. 2, pp. 193–234, 1983.

[8] Ryan A Rossi, Di Jin, Sungchul Kim, Nesreen K Ahmed, Danai Koutra, and John Boaz Lee, "On proximity and structural role-based embeddings in networks: Misconceptions, techniques, and applications," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 14, no. 5, pp. 1–37, 2020.

[9] Nataša Pržulj, "Biological network comparison using graphlet degree distribution," *Bioinformatics*, vol. 23, no. 2, pp. e177–e183, 2007.

[10] Ryan A Rossi, Rong Zhou, and Nesreen K Ahmed, "Estimation of graphlet counts in massive networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 1, pp. 44–57, 2018.

[11] Xutong Liu, Yu-Zhen Janice Chen, John CS Lui, and Konstantin Avrachenkov, "Learning to count: A deep learning framework for graphlet count estimation," *Network Science*, p. 30, 2020.

[12] Nesreen K Ahmed, Ryan A Rossi, John Boaz Lee, Theodore L Willke, Rong Zhou, Xiangnan Kong, and Hoda Eldardiry, "role2vec: Role-based network embeddings," in *Proc. DLG KDD*, 2019.

[13] Kathryn Cooper and Mauricio Barahona, "Role-based similarity in directed networks," *arXiv preprint arXiv:1012.2726*, 2010.

[14] K Henderson, B Gallagher, T Eliassi-Rad, H Tong, L Akoglu, D Koutra, L Li, S Basu, and C Faloutsos, "Rolx: Role extraction and mining in large networks," Tech. Rep., Lawrence Livermore National Lab, Livermore, CA (United States), 2011.

[15] Di Jin, Mark Heimann, Tara Safavi, Mengdi Wang, Wei Lee, Lindsay Snider, and Danai Koutra, "Smart roles: Inferring professional roles in email networks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2923–2933.

[16] Junchen Jin, Mark Heimann, Di Jin, and Danai Koutra, "Towards understanding and evaluating structural node embeddings," *arXiv preprint arXiv:2101.05730*, 2021.

[17] Arnaud Browet, *Algorithms for community and role detection in networks.*, Ph.D. thesis, Catholic University of Louvain, Belgium, 2014.

[18] Thomas P Cason, *Role extraction in networks.*, Ph.D. thesis, Catholic University of Louvain, Belgium, 2012.

[19] Santo Fortunato, "Community detection in graphs," *Physics reports*, vol. 486, no. 3-5, pp. 75–174, 2010.

[20] Leto Peel, Jean-Charles Delvenne, and Renaud Lambiotte, "Multiscale mixing patterns in networks," *Proceedings of the National Academy of Sciences*, vol. 115, no. 16, pp. 4057–4062, 2018.

[21] Leo Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.

[22] László Babai, "Graph isomorphism in quasipolynomial time," in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, 2016, pp. 684–697.

[23] Boris Weisfeiler and Andrei Leman, "The reduction of a graph to canonical form and the algebra which appears therein," *NTI, Series*, vol. 2, no. 9, pp. 12–16, 1968.

[24] László Babai and Ludik Kucera, "Canonical labelling of graphs in linear average time," in *20th Annual Symposium on Foundations of Computer Science*. IEEE, 1979, pp. 39–46.

[25] Jin-Yi Cai, Martin Fürer, and Neil Immerman, "An optimal lower bound on the number of variables for graph identification," *Combinatorica*, vol. 12, no. 4, pp. 389–410, 1992.

[26] Christopher Morris, Martin Ritzert, Matthias Fey, William L Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe, "Weisfeiler and leman go neural: Higher-order graph neural networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, pp. 4602–4609.

[27] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka, "How powerful are graph neural networks?," *arXiv preprint arXiv:1810.00826*, 2018.

[28] Asim Kumar Debnath, Rosa L Lopez de Compadre, Gargi Debnath, Alan J Shusterman, and Corwin Hansch, "Structure-activity relationship of mutagenic aromatic and heteroaromatic nitro compounds. correlation with molecular orbital energies and hydrophobicity," *Journal of medicinal chemistry*, vol. 34, no. 2, pp. 786–797, 1991.

[29] Ida Schomburg, Antje Chang, Christian Ebeling, Marion Gremse, Christian Heldt, Gregor Huhn, and Dietmar Schomburg, "Brenda, the enzyme database: updates and major new developments," *Nucleic acids research*, vol. 32, no. suppl_1, pp. D431–D433, 2004.

[30] Nikil Wale, Ian A Watson, and George Karypis, "Comparison of descriptor spaces for chemical compound retrieval and classification," *Knowledge and Information Systems*, vol. 14, no. 3, pp. 347–375, 2008.

[31] Christopher Morris, Nils M. Kriege, Franka Bause, Kristian Kersting, Petra Mutzel, and Marion Neumann, "Tudataset: A collection of benchmark datasets for learning with graphs," in *ICML 2020 Workshop on Graph Representation Learning and Beyond (GRL+ 2020)*, 2020.

[32] Thomas N Kipf and Max Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[33] Chen Cai and Yusu Wang, "A note on over-smoothing for graph neural networks," *arXiv preprint arXiv:2006.13318*, 2020.

[34] Kenta Oono and Taiji Suzuki, "Graph neural networks exponentially lose expressive power for node classification," *arXiv preprint arXiv:1905.10947*, 2019.

ANALYSIS OF CONTRACTIONS IN SYSTEM GRAPHS: APPLICATION TO STATE ESTIMATION

Mohammadreza Doostmohammadian[†⋆], Themistoklis Charalambous[†], Senior Member, IEEE, Miadreza Shafie-khah[*], Hamid R. Rabiee[◇], Senior Member, IEEE, and Usman A. Khan[‡], Senior Member, IEEE

† School of Electrical Engineering, Aalto University, Espoo, Finland.
⋆ Faculty of Mechanical Engineering, Semnan University, Semnan, Iran.
* School of Technology and Innovations, University of Vaasa, Vaasa, Finland
◇ Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
‡Electrical and Computer Engineering Department, Tufts University, Medford, MA, USA.

ABSTRACT

Observability and estimation are closely tied to the system structure, which can be visualized as a *system graph*–a graph that captures the inter-dependencies within the state variables. For example, in social system graphs such inter-dependencies represent the social interactions of different individuals. It was recently shown that contractions, a key concept from graph theory, in the system graph are critical to system observability, as (at least) one state measurement in every contraction is necessary for observability. Thus, the size and number of contractions are critical in recovering for loss of observability. In this paper, the correlation between the average-size/number of contractions and the global clustering coefficient (GCC) of the system graph is studied. Our empirical results show that estimating systems with high GCC requires fewer measurements, and in case of measurement failure, there are fewer possible options to find substitute measurement that recovers the system's observability. This is significant as by tuning the GCC, we can improve the observability properties of large-scale engineered networks, such as social networks and smart grid.

Index Terms— Contraction, clustering coefficient, structural observability, estimation, system graph

1. INTRODUCTION

Large-scale networked systems have seen a surge of interest in recent control and signal processing literature with applications in IoT and CPS [1–3]. A key challenge in such networks is state estimation [2, 4, 5] via a distributed network of measurements. From this perspective of distributed estimation, an effective tool is the system graph in which nodes represent state variables and edges between two nodes show coupling among the two state variables [6–8], motivating structural control and graph signal processing. In this sense, structural observability is related to certain system graph properties relying only on the system structure, and not on the exact system parameter values [4, 6, 7, 9].

An important graph-theoretic property to study system observability is the notion of *contraction* in the system graph, which is the dual of *dilation* in controllability [7]. In a contraction, multiple nodes are contracted (connected) to a fewer group of nodes. It is known that measuring one state node in every contraction is essential for network observability [10,11]. All states in a contraction are thus observationally equivalent which is significant for observability recovery, for example, in sensor/measurement failure [4, 11]. The size of contractions is also a key property, representing the number of possible options for estimation/observability recovery. A large contraction presents more choices of equivalent state measurements to replace the failed/faulty observation, or, for example, to minimize cost [12–14]. Also, the number of contractions represents the number of necessary measurement (or sensors) for estimation. The size and distribution of contractions in a system graph depend on certain graph properties. This work particularly studies how the *global clustering coefficient* affects the distribution of contractions. This paper is a nonlinear model extension of our previous works [1, 10] on *local clustering coefficient* and *degree heterogeneity*.

This paper models the nonlinear system as a random Scale-Free (SF) graph. The reason is that the structure of most real-world systems resemble the structure of SF graphs [15]. To study the effect of the GCC, as our main contribution, the distribution of size/number of contractions in SF graphs and clustered SF (CSF) graphs are compared. Due to specific formation in CSF graphs (known as *triad formation*) they have higher GCC, while their other properties (particularly

The work of U. Khan was supported by NSF under awards #1903972 and #1935555. The work of T. Charalambous was supported by the Academy of Finland under Grant 317726. Corresponding author email: doost@semnan.ac.ir, mohammadreza.doostmohamadian@aalto.fi.

power-law degree distribution) are similar to SF graphs. The significance of this contribution is that by tuning the network GCC, e.g., adopting the results of [16, 17], one can improve/impair system observability properties. Our results can be used in the design of large-scale man-made networks to improve their estimation properties in terms of reducing necessary observer nodes (sensor locations) for cost-optimal estimation. An example of such re-design of power grid is given in Section 5 as another contribution of this paper. Another possible application is in changing the structure of social networks to hinder the possibility of distributed estimation [18, 19] and, therefore, improve information privacy and reduce the vulnerability towards information leakage [20].

The rest of this paper is as follows. Section 2 describes graph-theory notions to define the contractions. Section 3 states the specific application to system estimation and observability. In Section 4, the distribution of contractions in SF and CSF graphs are compared, and an illustrative example application in power grid monitoring is given in Section 5. Conclusions and future research are presented in Section 6.

2. CONTRACTIONS IN GRAPHS

We consider the complex system, for example a social system, as an undirected graph or a strongly-connected directed graph (SC digraph) denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with the node set $\mathcal{V}$ (representing the n states) and edge/link set $\mathcal{E} = \{(v_i, v_j)\}$. The associated bipartite system graph $\Gamma = (\mathcal{V}^+, \mathcal{V}^-, \mathcal{E}_\Gamma)$ is defined with two disjoint left/right node sets $\mathcal{V}^+$ and $\mathcal{V}^-$, and edges $\mathcal{E}_\Gamma = \{(v_j^-, v_i^+) | (v_j, v_i) \in \mathcal{E}\}$. In $\mathcal{G}$, the edges with no common end node are called a matching $\underline{\mathcal{M}}$, which equivalently in Γ represent the subset of edges not incident on the same node in $\mathcal{V}^+$. In other words, $\underline{\mathcal{M}}$ is a set of pairwise disjoint edges (with no loop). A matching with maximum size is called maximum (cardinality) matching $\mathcal{M}$, which is not a subset of any other matching. Note that there are many possible choices of $\mathcal{M}$ in general. The nodes respectively in $\mathcal{V}^+$ and $\mathcal{V}^-$ incident to the chosen $\mathcal{M}$ are denoted by $\partial\mathcal{M}^+$ and $\partial\mathcal{M}^-$, and the nodes in $\delta\mathcal{M} = \mathcal{V}^+ \backslash \partial\mathcal{M}^+$ are unmatched in $\mathcal{V}^+$. Given $\mathcal{M}$, let $\Gamma^{\mathcal{M}}$ be the auxiliary graph made by reversing all edges in $\mathcal{M}$, and holding all the other edges $\mathcal{E}_\Gamma \backslash \mathcal{M}$ in Γ. In $\Gamma^{\mathcal{M}}$, an alternating path associated to $\mathcal{M}$ (also called $\mathcal{M}$-alternating path), denoted by $\mathcal{Q}_\mathcal{M}$, is a path starting from a node in $\delta\mathcal{M}$ with its edges alternately in $\mathcal{M}$ and not in $\mathcal{M}$. An augmenting path $\mathcal{P}_\mathcal{M}$ associated to $\mathcal{M}$ (also called $\mathcal{M}$-augmenting path) is an $\mathcal{M}$-alternating path in $\Gamma^{\mathcal{M}}$ that starts from and ends in $\delta\mathcal{M}$. For a matching $\underline{\mathcal{M}}$ and associated $\mathcal{P}_{\underline{\mathcal{M}}}$, $\underline{\mathcal{M}} \oplus \mathcal{P}_{\underline{\mathcal{M}}}$ represents a new matching with one more edge than $\underline{\mathcal{M}}$, where $\oplus$ is the XOR operator. *In $\Gamma^{\mathcal{M}}$, a contraction $\mathcal{C}_j$, associated to an unmatched node $v_j \in \delta\mathcal{M}$, is defined as the set of all state nodes in $\mathcal{V}^+$ reachable by $\mathcal{M}$-alternating paths starting from v_j.* Intuitively speaking, contraction represents subset of nodes linking to smaller subset of nodes [10,21]. Algorithm 1 [10,21] presents the pseudo-code for finding graph contractions with polynomial order complexity $\mathcal{O}(n^{2.5})$. Polynomial complexity facilitates applications in large-scale as in social networks or power grids.

Algorithm 1: Finding contractions in a graph [10, 21].

Given: System graph $\mathcal{G}$
Find Γ;
Find $\underline{\mathcal{M}}$;
Find $\Gamma^{\underline{\mathcal{M}}}$;
while $\mathcal{P}_{\underline{\mathcal{M}}}$ *exist* **do**
  for *nodes in* $\delta\underline{\mathcal{M}}$ **do**
    Find $\mathcal{P}_{\underline{\mathcal{M}}}$;
    $\underline{\mathcal{M}} = \underline{\mathcal{M}} \oplus \mathcal{P}_{\underline{\mathcal{M}}}$;
Find $\Gamma^{\mathcal{M}}$;
for *state nodes in* $\delta\mathcal{M}$ **do**
  Find $\mathcal{Q}_\mathcal{M}$ in $\Gamma^{\mathcal{M}}$;
  Put nodes in $\mathcal{V}^+$ reachable by $\mathcal{Q}_\mathcal{M}$ in $\mathcal{C}_i$;
Return $\mathcal{C}_i, i = \{1, ..., l\}$;

In this paper, as our main contribution, we aim to understand possible correlation between the GCC and prevalence of contractions, and interpret the implication of this relation through a system estimation perspective.

3. APPLICATION TO STATE ESTIMATION

In this work, a *nonlinear* autonomous dynamic system (in contrast to the linear model in [10]) is considered as,

$$\dot{\mathbf{x}} = f(\mathbf{x}(t)) + \mathbf{v}, \tag{1}$$

where the state variable $\mathbf{x} = [x^1, \ldots, x^n]^\top \in \mathbb{R}^n$ is to be estimated via the measurements,

$$\mathbf{y}(t) = g(\mathbf{x}(t)) + \mathbf{r}, \tag{2}$$

where $\mathbf{y} = [y^1, \ldots, y^m] \in \mathbb{R}^m$ is the measurement, and $\mathbf{v}$ and $\mathbf{r}$ are Gaussian noise. The system model (1)-(2) can be represented as a Linear-Structure-Invariant (LSI) model as,

$$\dot{\mathbf{x}} = A(t)\mathbf{x}(t) + \mathbf{v}, \tag{3}$$
$$\mathbf{y}(t) = C(t)\mathbf{x}(t) + \mathbf{r}, \tag{4}$$

where $A(t)$ and $C(t)$ are time-dependent system and measurement matrices representing the linearization of the system and measurement functions $f(\cdot)$ and $g(\cdot)$ over time. Recall that, from Kalman filtering theory, the underlying system can be estimated if it is *observable* via the given measurements. System observability implies that the global vector, $\mathbf{x}$, can be uniquely determined by the measurements, $\mathbf{y}$. As shown in [7], the observability of the nonlinear model (1)-(2) is equivalent with the observability of the linearized model (3)-(4) over all operating points. The structure (the zero-nonzero pattern) of the associated linearized matrices

$A(t)$ and $C(t)$ are time-invariant while the numerical values of their nonzero entries may vary at different operating points, implying the name structure-invariant. This motivates the concept of *structural observability* (or *generic observability*) based on structured systems theory [6, 7, 9], which provides a graph-theoretic method to check for system observability. In structural analysis the system is modeled as a *system graph*, where a node v_i models a state x^i and a link $v_j \to v_i$ models the dependency of the two state variables x^i and x^j. In other words, if f^i is a function of x^j then the entry $\frac{\partial f^i}{\partial x^j}$ in linearized matrix A is nonzero while its exact value depends on the operating point and may change over time [9, 19]. Denote the system graph by $\mathcal{G}_A = (\mathcal{V}, \mathcal{E}_A)$, with state nodes $\mathcal{V}$ and $\mathcal{E}_A = \{(v_j, v_i) \mid \frac{\partial f^i}{\partial x^j} \neq 0\}$ including the edges $v_j \to v_i$. Using the definitions in Section 2, the next theorem states necessary conditions for observability of the system graph $\mathcal{G}_A$.

Theorem 1 *Let $\delta\mathcal{M}$ denote the set of unmatched nodes of system graph $\mathcal{G}_A$ associated with an autonomous LSI system. To ensure observability, it is necessary to measure every unmatched state in $\delta\mathcal{M}$.*

We refer to [7] for the proof (in the dual case of controllability). Based on the definition, for a given maximum matching $\mathcal{M}$, every node $v_j \in \delta\mathcal{M}$ belongs to a contraction $\mathcal{C}_i$, while the nodes $\mathcal{C}_i \backslash v_j$ are all matched. This leads to the following *observational equivalence* property in contractions:

Theorem 2 *Consider an LSI system abstracted as a graph (undirected or SC) with contractions $\mathcal{C} = \{\mathcal{C}_1, ..., \mathcal{C}_m\}$. The necessary condition for observability is to measure (at least) one state node in every contraction.*

Corollary 1 *Nodes in a contraction are equivalent for observability recovery.*

See the proofs in [10, 11]. This corollary implies that when a critical measurement fails, causing loss of observability, measurement of any other state node in the same contraction recovers the system observability [4]. Recall that (structural) rank deficiency of the system matrix A defines the cardinality of $\delta\mathcal{M}$ and $\mathcal{C}$ in its associated system graph $\mathcal{G}_A$ [21, 22].

4. CONTRACTION PREVALENCE IN SF GRAPHS

In this section, the distribution of contractions in SF networks, as random graph-representations of real-world systems, is studied. Such random models simplify the understanding of different processes, e.g., spreading processes and cascading failures [15, 23–25]. The main feature of SF graphs is their *power-law degree distribution* [15], which implies that there are few hubs (nodes with high degree) and large number of low-degree nodes in the SF network. To construct such networks, Ref. [15] provides an iterative algorithm initializing with a small *seed graph* and recursively adding a new node with m new edges. The main feature of this iterative procedure is that the linking probability between the new node and an old node is proportional to its degree. In other words, the new node *prefers* connecting to old hubs, hence it is named *preferential attachment*. Such SF graphs are known to have low GCC[1], while real-world networks, for example social networks, show high clustering. Therefore, a modified model with high GCC is proposed [23–25], named Clustered Scale Free (CSF). The building blocks of this model are similar to the preferential attachment, where the difference is the *triad formation* step. In this model, the newly added node directly links to m_r nodes, while also making m_s preferential linking to some neighbors of m_r preferentially attached nodes to create triads. This significantly increases the GCC in CSF networks with the same average node degree as SF networks.

4.1. Empirical results and simulation

To study the effect of GCC, we compare the number/average-size of contractions in CSF/SF graphs. We perform Monte-Carlo simulations over 50 realizations of sample CSF and SF graphs with $m = m_r + m_s = 2$ and $n = 100$ to $n = 1000$ nodes. Having $m = m_r + m_s$ ensures equal number of new edges via preferential attachment in both CSF and SF networks, implying the same average node degrees for both types. This is essential for comparison as all features of both CSF and SF graphs must be similar while only their GCC differs [23]. Fig. 1 shows the Monte-Carlo simulation results.

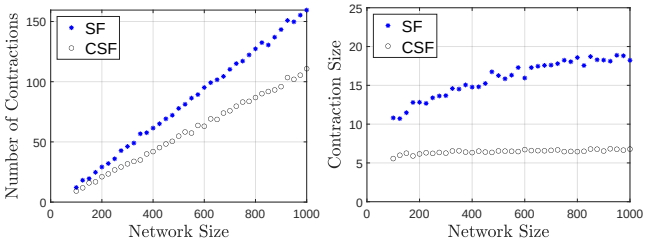


Fig. 1: Number and size of contractions averaged over 50 realizations of growing SF and CSF networks.

As shown in Fig. 1, in SF graphs there are more contractions which are in average larger (in size) as compared to CSF graphs. Table 1 summarizes the results shown in Fig. 1.

Table 1: Comparison of CSF and SF graph samples.

Graph Type	SF	CSF
Average size of Contraction	15.8	6.6
Average number of Contractions	86	59
Average GCC	0.017	0.181

[1]GCC is defined as the ratio of the triangles tr to the total number of connected open triplets trp in the graph, i.e., GCC $= \frac{3.tr}{trp}$ [15].

4.2. Discussions on the results

From Fig. 1 and Table 1, we see the average size of contractions in SF networks is significantly larger than CSF networks. Recall that both graph types are constructed via the preferential attachment method and, therefore, both share similarity of most graph properties, e.g., (i) logarithmic growth in mean shortest-path and (ii) power-law degree distribution [23]. Their main difference stems from the GCC, which is lower in SF graphs. This is the key feature contributing to the decrease in both size and number of contractions in CSF graphs. Again we emphasize that the other graph properties of both types are similar. Thus, one can conclude that, in graphs with power-law degree distribution, increase in GCC causes fewer contractions with smaller average-size.

In terms of system observability/estimation, the implication is that for graphs with higher GCC: (i) fewer state measurements are needed to ensure observability; and (ii) fewer observationally equivalent states are available to recover observability loss in case of measurement failure. The first result deduced from number of contractions while the latter stems from average size of contractions. This implies that the observability (and consequently estimation properties) can be improved/deteriorated by tuning the GCC of (synthetic) system networks via [1, 16, 17]. A such example is given next.

5. ILLUSTRATIVE EXAMPLE: APPLICATION TO POWER NETWORKS

The power grid can be conceived as a large-scale dynamical system [26], where the sparsity of its system matrix follows the structure of the distribution network [27]. To ensure reliable power delivery, the electrical phasor states (voltage, current, etc.) are typically measured via phasor measurement units (PMUs), distributed over the electricity grid. The PMU placement is such that to ensure observability of the entire power grid for monitoring purposes [28]. From Section 4, one can reduce the number of allocated PMUs for observability by re-designing the power network structure. Consider the European power grid [29] shown in Fig. 2(top) with 1494 nodes (buses) and 2156 edges (transmission lines). The unmatched nodes represent the possible PMU placements in the grid. We increase the GCC by adding 40 edges between certain bus-nodes as shown in Fig. 2(bottom). The network grid properties before and after edge addition are compared in Table 2. Note that the change in average node degree is negligible while the GCC is increased by 19%. As expected, unmatched nodes (contractions) are reduced by 47% via adding only 40 edges (about 1.8% of the total edges). The motivation behind this example is to show that by design of power networks with higher GCC, the number of PMU placements can be reduced while having the same number of edges (transmission lines) as discussed in Section 4.2. This implies lower monitoring cost with the same cost for infrastructural network.

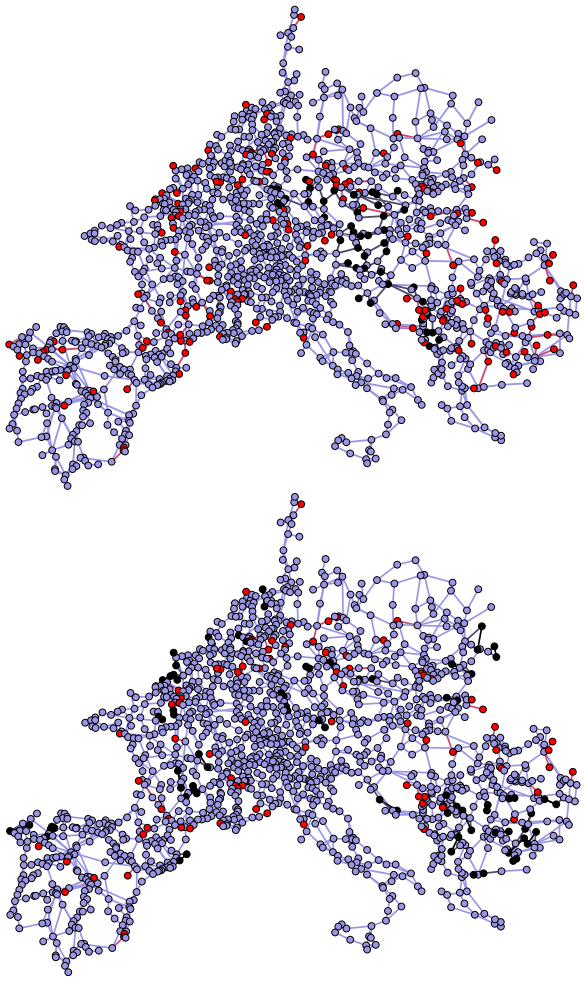


Fig. 2: (top) The European power grid with 151 unmatched bus-nodes (shown in red) whose states need to be measured via PMUs. An example contraction of 46 nodes is also shown in black. (bottom) By adding 40 edges (shown in black) to increase the GCC of power network, the number of unmatched bus-nodes (and minimum required PMUs) is reduced to 80.

Table 2: Network characteristics before/after link addition.

links	average degree	GCC	contractions
2156	2.886	0.094	151
2196	2.939	0.112	80

6. CONCLUSIONS AND FUTURE RESEARCH

In this work, distribution of contractions in SC digraphs and its relation with graph GCC is studied. Note that for general non-SC digraphs, another component known as *root SCC* or *parent SCC* also affects system graph observability [30, 31]. As future research, we intend to study the correlation between prevalence/size of both parent SCCs and contractions with other graph features, such as assortativity/disassortativity and community/hierarchical structure [15].

7. REFERENCES

[1] M. Doostmohammadian and H. R. Rabiee, "On the observability and controllability of large-scale iot networks: Reducing number of unmatched nodes via link addition," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1747–1752, 2020.

[2] M. Rana, L. Li, and S. W. Su, "Distributed state estimation over unreliable communication networks with an application to smart grids," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 1, pp. 89–96, 2017.

[3] H. Arasteh, V. Hosseinnezhad, V. Loia, A. Tommasetti, O. Troisi, M. Shafie-khah, and P. Siano, "Iot-based smart cities: A survey," in *IEEE 16th International Conference on Environment and Electrical Engineering*. IEEE, 2016, pp. 1–6.

[4] M. Doostmohammadian, H. R. Rabiee, H. Zarrabi, and U. A. Khan, "Distributed estimation recovery under sensor failure," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1532–1536, 2017.

[5] M. Kabiri, N. Amjady, M. Shafie-khah, and J. P. S. Catalao, "Enhancing power system state estimation by incorporating equality constraints of voltage dependent loads and zero injections," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 659–671, 2018.

[6] J. M. Dion, C. Commault, and J. van der Woude, "Generic properties and control of linear structured systems: A survey," *Automatica*, vol. 39, pp. 1125–1144, Mar. 2003.

[7] Y. Y. Liu, J. J. Slotine, and A. L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, May 2011.

[8] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[9] J. F. Carvalho, S. Pequito, A. P. Aguiar, S. Kar, and K. H. Johansson, "Composability and controllability of structural linear time-invariant systems: Distributed verification," *Automatica*, vol. 78, pp. 123–134, 2017.

[10] M. Doostmohammadian, H. R. Rabiee, H. Zarrabi, and U. Khan, "Observational equivalence in system estimation: Contractions in complex networks," *IEEE Transactions on Network Science and Engineering*, vol. 5, no. 3, pp. 212–224, 2017.

[11] C. Commault, J. Dion, D. H. Trinh, and T. H. Do, "Sensor classification for the fault detection and isolation, a structural approach," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 1, pp. 1–17, 2011.

[12] B. Guo, O. Karaca, T. H. Summers, and M. Kamgarpour, "Actuator placement under structural controllability using forward and reverse greedy algorithms," *IEEE Transactions on Automatic Control*, 2020.

[13] S. Moothedath, P. Chaporkar, and M. N. Belur, "Minimum cost feedback selection for arbitrary pole placement in structured systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3881–3888, 2018.

[14] M. Doostmohammadian, H. R. Rabiee, and U. A. Khan, "Structural cost-optimal design of sensor networks for distributed estimation," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 793–797, 2018.

[15] M. Newman, A. Barabási, and D. J. Watts, *The structure and dynamics of networks.*, Princeton university press, 2006.

[16] N. Islam, "Towards a secure and energy efficient wireless sensor network using blockchain and a novel clustering approach," M.S. thesis, Dalhousie University, 2018.

[17] J. M. Moore, M. Small, and G. Yan, "Inclusivity enhances robustness and efficiency of social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 563, pp. 125490, 2021.

[18] E. Montijano, G. Oliva, and A. Gasparri, "Distributed estimation and control of node centrality in undirected asymmetric networks," *IEEE Transactions on Automatic Control*, 2020.

[19] M. Doostmohammadian, H. R. Rabiee, and U. A. Khan, "Cyber-social systems: modeling, inference, and optimal design," *IEEE Systems Journal*, vol. 14, no. 1, pp. 73–83, 2019.

[20] J. Lovato, A. Allard, R. Harp, and L. Hébert-Dufresne, "Distributed consent and its impact on privacy and observability in social networks," *arXiv preprint arXiv:2006.16140*, 2020.

[21] K. Murota, *Matrices and matroids for systems analysis*, Springer, 2000.

[22] M. Doostmohammadian and U. A. Khan, "On the distributed estimation of rank-deficient dynamical systems: A generic approach," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 4618–4622.

[23] P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering," *Physical review E*, vol. 65, no. 2, pp. 026107, 2002.

[24] I. Türker, "Generating clustered scale-free networks using poisson based localization of edges," *Physica A: Statistical Mechanics and its Applications*, vol. 497, pp. 72–85, 2018.

[25] C. P. Herrero, "Ising model in clustered scale-free networks," *Physical Review E*, vol. 91, no. 5, pp. 052812, 2015.

[26] J. Machowski, Z. Lubosny, J. W. Bialek, and J. R. Bumby, *Power system dynamics: stability and control*, John Wiley & Sons, 2020.

[27] U. A. Khan and M. Doostmohammadian, "A sensor placement and network design paradigm for future smart grids," in *4th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Puerto Rico, 2011, pp. 137–140.

[28] R. Babu and B. Bhattacharyya, "Optimal allocation of phasor measurement unit for full observability of the connected power network," *International Journal of Electrical Power & Energy Systems*, vol. 79, pp. 89–97, 2016.

[29] "The RE-Europe data set," https://zenodo.org/record/35177.

[30] S. Pequito, V. M. Preciado, A. Barabási, and G. J. Pappas, "Trade-offs between driving nodes and time-to-control in complex networks," *Scientific reports*, vol. 7, no. 1, pp. 1–14, 2017.

[31] M. Doostmohammadian and U. A. Khan, "On the characterization of distributed observability from first principles," in *2nd IEEE Global Conference on Signal and Information Processing*, 2014, pp. 914–917.

FAST MACHINE LEARNING-BASED SIGNAL CLASSIFICATION IN ENERGY CONSTRAINED CRN: FPGA DESIGN AND IMPLEMENTATION

Arash Rasti-Meymandi[1], *Jamshid Abouei*[1], *Zohreh Hajiakhondi-Meybodi*[2], *Arash Mohammadi*[3], *Amir Asif*[4]

[1]Department of Electrical Engineering, Yazd University, Yazd, Iran
[2]Department of Electrical Engineering, Concordia University, Montreal, Canada
[3]Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, Canada
[4]Department of Electrical Engineering, York University, Toronto, Canada

ABSTRACT

Cognitive Radio Networks (CRNs) is positioned as an appealing autonomous system to enhance spectrum scarcity by dynamic spectrum access and spectrum sharing across wireless networks. To operate at the highest performance level, the allocation and vacation process of primary and secondary users need to be accomplished rapidly. This issue motivates us to propose a fast machine learning-based processing algorithm, referred to as the Arithmetic Shifter-Based Support Vector Machine (ASB-SVM) classifier. The novelty of our proposed scheme is to increase the speed of signal classification by employing shift registers in a two multipliers feature mapping method instead of using multiplication blocks in the SVM classifier. The proposed ASB-SVM design is implemented in Xilinx Virtex-6 XC6VLX240T FPGA. By exploiting spectral features for the classifier, an overall accuracy rate of 98.2% is achieved for green modulated signals in CRNs. Experimental results show that given the feature vector, our proposed system is capable of classifying a blind modulated signal within just 3 ns in the classifier block of a CRN while achieving 30% resource reduction and 45% increase in speed compared to the conventional linear SVM implementation.

Index Terms— Cognitive Radio, Primary User Signal Detection, Machine Learning, FPGA

1. INTRODUCTION

Cognitive Radio Networks (CRNs) have been proven to be the most promising approach towards efficient spectrum utilization in wireless communication industry [1]. Their main goal is to create autonomous systems that continuously monitor the spectrum, decide, and act upon accordingly [2]. Autonomous systems are referred to intelligent systems capable of operating without human intervention [3]. They are unltimately aiming to implement brain-inspired systems capable of operating with a human counterpart without imperative instructions [4].

The most critical and time-consuming part of CRNs is the spectrum sensing process, in which the shared spectrum is used by an unlicensed Secondary User (SU) as long as its imposed interference on the receiver of the licensed Primary User (PU) is below an acceptable level. The absence of PUs is necessary to assign certain white spaces to SUs. Nevertheless, once PUs launch their transmissions, SUs should vacate the spectrum. The spectrum allocation and vacation cycle must be completed quickly so that CRNs can operate at the highest performance and productivity levels, otherwise an interference between the PU and SU occurs. In fact, the quick detection of the PU is critical since violation of the detection on time will cause an interference to the PU. However, due to the spectrum shortage and transmission power constraint of mobile devices, recent CRNs have been recognized as green communication technologies [5]. In this regard, the main objective of spectrum sensing at energy limited SUs is to detect and identify the signal of the PU in a noisy environment with a high precision rate and as quick as possible [6]. This issue motivates researchers to come up with the fast signal processing algorithms to achieve an easy software implementation with the focus on efficient signal classification schemes. One of the promising platforms for such classification scheme implementation is the FPGAs. These reconfigurable hardware devices provide suitable infrastructures to embed autonomous systems. However, one downside to utilizing FPGAs is their large reconfiguration time [7]. Nevertheless, since PUs' signal modulations in CRNs are rarely prone to change with respect to the reconfiguration time of FPGAs, the slow reconfiguration will not cause serious problem to the whole system.

Literature Review: Spectrum sensing for signal classification is categorized into transmitter-based and machine learning-based approaches. In the transmitter-based case, the main goal is to detect the weak signal transmitted by PUs. Considering the fact that the transmitter-based spectrum sensing method depends on the noise level and prior knowledge of PU signals, it suffers from high complexity and computational capacity. Therefore, the focus of recent research works [8, 9, 10] have been shifted to machine learning-based models.

Reference [11] investigated the use of Genetic Programming (GP) in combination with K-Nearest Neighbor (KNN) for modulated signal classification. The authors in [12, 13] employed deep learning to classify modulated signals in low SNRs in CRNs. They achieved a high degree of accuracy to differentiate the signals, however, the complexity of resource utilization of such a system is high due to the use of convolutional neural network, leading to a reduction in the speed of signal classification process. The authors in [14, 15, 16] used deep learning to classify different modulations using spectral correlation function for extracting features. Since the spectrum sensing process needs to be performed fast, there have been some studies on the spectrum sensing implementation in FPGA. The authors in [17] implemented a cyclostationary feature detector only for Orthogonal Frequency Division Multiplexing (OFDM) signals in CRNs. In [18], an automatic modulation recognition implementation was introduced using spectral features. They achieved an overall accuracy rate of 90% in signal classification which might not be an acceptable rate in CRNs. The authors in [19] designed an FPGA IP core to speed up the classification process in CRNs. They utilized the Support Vector Machine (SVM) to effectively implement a high speed classification algorithm. This scheme, however, is only applicable for specific digital modulations that use complex symbols from a constellation.

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

Contribution: Increasing the performance of spectrum sensing in terms of speed and resource efficiency motivated us to design a novel machine learning-based processing algorithm that boosts the signal classification process with the minimum accuracy rate of 98.2%. The proposed scheme in conjunction with linear SVM is designed and implemented in Xilinx Virtex-6 XC6VLX240T FPGA. The designed algorithm benefits from the advantage of employing only shift registers instead of multiplication blocks in the SVM classifier that significantly enhances the speed of signal classification. Experimental results show that given the feature vector, our proposed approach is capable of classifying a blind modulated signal within only 3 ns in the classifier block in CRNs which corresponds to over 300 million signals per seconds throughput. In contrast to conventional schemes (e.g., [17, 19]), our proposed algorithm is independent of type of modulation, therefore, it can be applied to variety of signal classification process.

2. SYSTEM MODEL AND PROBLEM DESCRIPTION

We consider a CRN consisting of K primary users indexed by $PU_1, \ldots, PU_K$ and M secondary users denoted by $SU_1, \ldots, SU_M$. For indicating the transmitter and receiver pair at PUs, we use notation (T_{PU_j}, R_{PU_j}), and (T_{SU_i}, R_{SU_i}) to represent the opportunistic SUs transmitter and receiver pair. We assume both SU_i and PU_j operate in the same frequency band restricted by the CRN regulation. Indeed, in the absence of PU_j, opportunistic SUs are allowed to fill the shared spectrum [20]. However, the active SUs must vacate the spectrum once the PU_j starts the transmission, hence, the main task of each R_{SU_i} is to recognize T_{PU_j}'s signal from other T_{SU_k}'s signals, $k \neq j$, to occupy the spectrum hole [21]. In addition, various green modulation schemes, including M-PSK, M-FSK, and M-ASK, $M = 2, 4, \ldots$ are chosen for transmission and signal classification.

We employ five spectral features namely (i) the maximum value of the spectral power density of the normalized-centered instantaneous amplitude, (ii) the standard deviation of the absolute value of the centered non-linear component of the instantaneous phase, (iii) the standard deviation of the centered non-linear component of the direct instantaneous phase, (iv) the standard deviation of the absolute value of the normalized-centered instantaneous amplitude, and (v) the standard deviation of the absolute value of the normalized-centered instantaneous frequency [22]. The segment size which is the number of samples from which features are extracted is set to $Seg = 2048$. To include non-linearity to the hyperplane, a set of non-linear selective features is evaluated from the original five features (similar to applying non-linearity to logistic regression classifier) to construct our feature space denoted by $\mathbb{R}^d$, where $d = 11$. To avoid interference in CRN, the signal classification process needs to be quick. Toward this goal, one approach is to use the SVM classifier which is one of the mostly used algorithms in classification. SVM is capable of finding the best hyperplane with sufficient margin to discriminate against each class. Despite most recent literature (e.g., [23]), we apply linear kernel in our work since hyperplane parameters are easily extracted and stored for constructing the decision function. Moreover, linear kernels have low computational cost in comparison to the non-linear ones, and also they show a much better performance when adopted with our feature mapping method.

Suppose there are L modulations for classification which are known for all SUs. It is also assumed that the dataset numbers are equal for all classes. The training dataset is indicated as $T^{(t)} = \{(X_1^{(t)}, Y_1^{(t)}), \ldots, (X_N^{(t)}, Y_N^{(t)})\}$, $t = 1, \ldots, M$, where N is the number of training set, $X_i^{(t)} \in \mathbb{R}^d$, and $Y_j^{(t)} \in \{-1, 1\}$. Since our case is a multiclass classification, the *one-vs-one* method is used. For our L modulation class, $\frac{L(L-1)}{2}$ classifiers are constructed. The class of an unknown signal is done according to the maximum voting scheme, where each classifier votes for one class. Assuming two arbitrary classes, the SVM objective is to find the decision function $d(X) \in \mathbb{R}^d$ by calculating the SVM Wolf's dual problem given as

$$\max_{\alpha_n}\{\sum_{n=1}^{N}\alpha_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}\alpha_n\alpha_m Y_n^{(t)}Y_m^{(t)}k(X_n^{(t)}, X_m^{(t)})\},$$

$$s.t. \quad \sum_{n=1}^{N}\alpha_n Y_n^{(t)} = 0, \qquad 0 < \alpha_n < C, \tag{1}$$

where α_n is the Lagrangian multipliers and $k(X_n^{(t)}, X_m^{(t)})$ is the kernel function produced by the dot multiplying of mapping functions represented as $\langle\phi(X_n^{(t)}).\phi(X_m^{(t)})\rangle$. By applying linear kernel $k(X_n^{(t)}, X_m^{(t)}) = \langle X_n^{(t)}.X_m^{(t)}\rangle$ and solving (1) using quadratic programming, the parameters of the hyperplane for constructing the decision function is derived as $W^{(t)} = \sum_{\alpha_n>0}\alpha_n Y_n^{(t)}X_n^{(t)}$ and $w_0^{(t)} = Y_s^{(t)} - \sum_{\alpha_n>0}\alpha_n Y_n^{(t)}X_n^{(t)^T}X_s^{(t)}$, where (s) indicates a support vector sample. Hence, the decision function $d(X)$ for classifying a new feature vector is represented by

$$d(X) = \text{sign}(w_0^{(t)} + W^{(t)^T}X^{(t)}). \tag{2}$$

We will further use (2) and the derived hyperparameters and revise them to construct a faster efficient decision function suitable for FPGA implementation.

3. PROPOSED SIGNAL CLASSIFICATION

One of the most time consuming parts of machine learning algorithms such as linear SVM is the multiplication of their parameters with the feature (input) vector. In this paper, we design an Arithmetic Shifter-Based SVM (ASB-SVM), in which the multiplications is replaced by only shifters with the use of a new preprocessing method. ASB-SVM speeds up the classification process with a negligible reduction in the overall accuracy rate of the system performance. Fig. 1 illustrates the block diagram of the designed scheme for signal classification located at each SU's receiver. In the proposed algorithm, the classifier's input has been multiplexed with multiple preprocessed blind modulated signals. This will help take advantage of our designed fast performance classifier as will be demonstrated further throughout the paper.

3.1. Proposed two multipliers feature mapping method

There are two popular types of feature scaling: $i)$ standardization (zero-score normalization) where the feature space is rescaled so that each feature vector has the mean value of 0 and standard deviation of 1, and $ii)$ Min-Max normalization where the feature data is mapped into any arbitrary scale. Feature scaling will help learning algorithms to work more efficiently. It is common to scale the data between 0 and 1 or -1 and 1 which can be accomplished by

$$x_i^s = a + \frac{(x_i - \min(X_i))(b-a)}{\max(X_i) - \min(X_i)}, \tag{3}$$

where $i = 1, \ldots, d$, a and b denote minimum and maximum values of the scale, x_i indicates i^{th} feature value and X_i is i^{th} feature vector. The objective of our two multipliers feature mapping method is to map data in the power of two. The main reason for this mapping procedure is that it is possible to replace the multiplication of input and parameter vector by a shift register which is performed much faster than multiplication and uses fewer energy. In

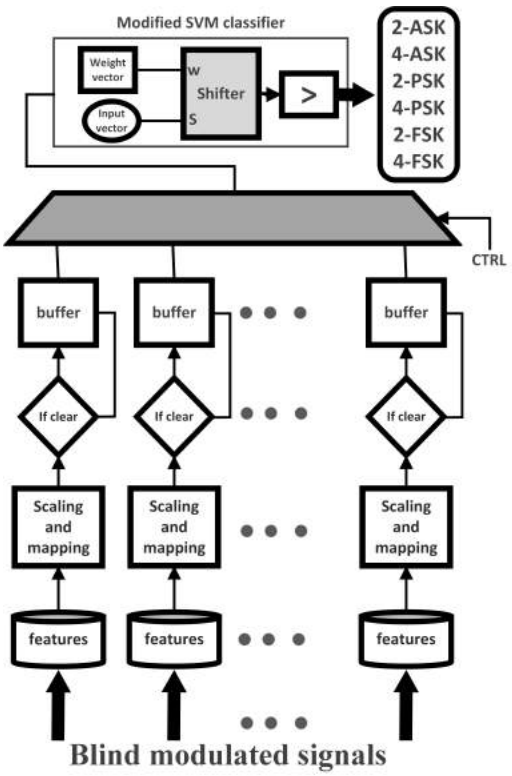


Fig. 1. Block diagram of the proposed system for signal classification located at each SU's receiver.

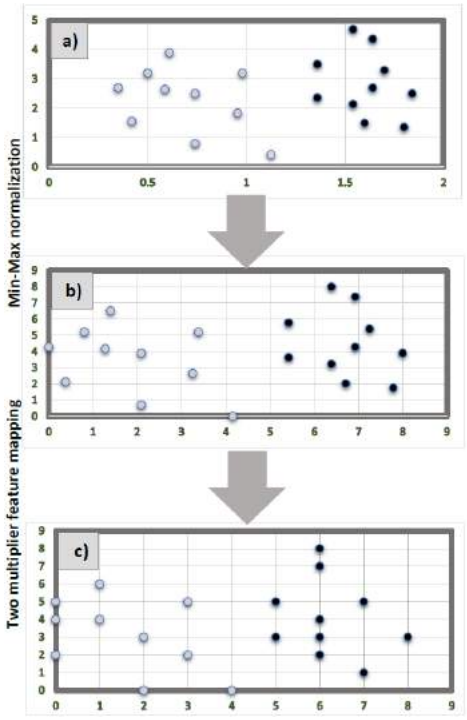


Fig. 2. The two multipliers feature mapping process for two arbitrary classes (the third figure from above has base-2 logarithmic scale).

this scheme, data is first scaled between 0 and 10 using (3). This helps data to be scattered through the feature space. The scaled data is mapped to the smallest near integer value using floor function $\lfloor \cdot \rfloor$.

Algorithm 1 ASB-SVM signal classification at SU's receiver

Input: $X = [x_1, \ldots, x_5]$, the hyperparamtes $\widehat{W} = [W_1, \ldots, W_{15}]$ for classifiers where $W_j = [w_0, \ldots, w_d]$, control signal c, and $CTRL$

Initialization: Clear all buffers B_i, $CTRL = 0$, $c = 1$.

$X^{new} \leftarrow$ Built new features from X.
$X^s \leftarrow$ Scale X^{new} to (0–10) by Min-Max normalization.
$B_i \leftarrow X^{map} \leftarrow$ Map X^s to smallest near value using $\lfloor \cdot \rfloor$.
while c **do**
  if B_i is ready **then**
    for $j = 1$ to number of classifier's block **do**
      $d(x)^j_{new} \leftarrow \text{sign}(W_j <<< X^{map})$
    end for
    use $d(x)^j_{new}$ values to classify the signal using
    maximum voting scheme
  else
    Go to next buffer uisng CTRL
  end if
  Output: Natural value of n for one of the six modulation.
end while

This will create a new feature space denoted by $\mathbb{Z}^d$. The final stage of this process is accomplished through converting the values of feature space $\mathbb{Z}^d$ to power of two, leading to creation of new feature space $\widehat{\mathbb{Z}}^d$. The visualization of the scheme is illustrated in Fig. 2. Hence, the mapping process can be expressed as $\widehat{X}^s = 2^{\lfloor X^s \rfloor}$, where $X^s = [x^s_1, \ldots, x^s_d]$. Training the classifier with the new feature space $\widehat{\mathbb{Z}}^d$ gives the opportunity to substitute multiplications with only shifters since shifting the parameter vector W with X^s is the same as multiplying W with $\widehat{X}^s$ (e.g. $8 \times w = 2^3 \times w \equiv (w$ arithmetic right shift by 3)). The same procedure in Section 2 for training the Multiplication-Based SVM (MB-SVM) classification is applied to obtain the hyperparameters of the ASB-SVM classifier using new training set $(\widehat{X^s_i}, Y_j)$. Due to the fact that the input vector is in the order of power two, the decision function in (2) can be rewritten as

$$d(X)_{new} = \text{sign}(\widehat{W} <<< X^{map}), \tag{4}$$

where $X^{map} = [\lfloor 0, x^s_1, \ldots, x^s_d \rfloor]$ denotes the feature vector, $\widehat{W} = [w_0, W]$ indicates the hyperparameter vector, and the notation $\langle <<< \rangle$ is the elementwise arithmetic shift. Algorithm 1 summarizes the whole procedure of our proposed system for signal classification. Note that the designed classifier specialized for FPGA implementation is trained in a supervised manner. It autonomously performs the signal classification once it is fully trained. Therefore, if the classifier is required to adapt to a new context such as a new signal modulation, its corresponding training data will still be required. this constraint, however, does not culminate in a challenging task since signal modulations are synthetic and can be readily generated and manipulated to procure the new training data.

Remark 1: One possible problem that can be emerged with this mapping technique is the overlapping of different class data in the new feature space $\mathbb{Z}^d$. To address this issue, we numerically tested whether any mapped sample has overlapped with another sample from different classes. The result showed that the more feature space we have, the less it is possible for the mapped sample to be overlapped with a sample from another class. Moreover, the problem can be avoided if classes are separable with enough margin. In the case of the aforementioned modulations for classification, there was not any overlapping of different sample classes as the number of our features is sufficient enough and the data set classes have a tolerable

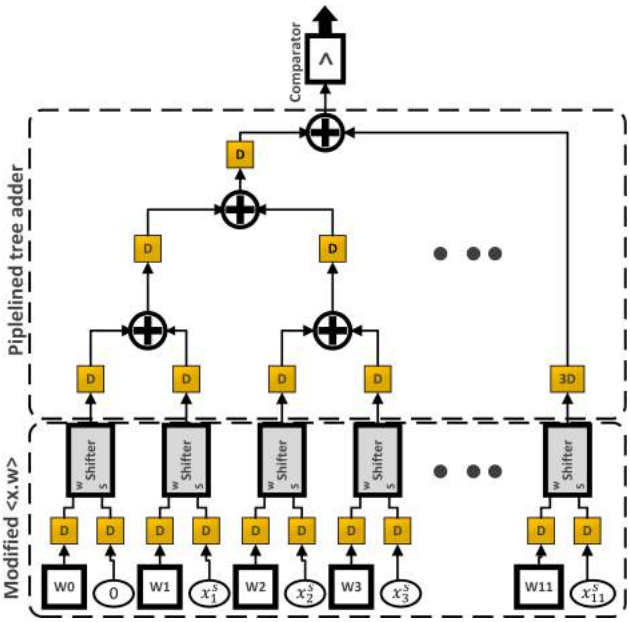


Fig. 3. One out of fifteen blocks of the ASB-SVM classifier.

margin in feature space $\mathbb{R}^d$.

Remark 2: It should be noted that there is another approach for effectively applying digital multipliers, called Canonical Signed Digit (CSD) multipliers. In the CSD method, a constant value is converted in such a way that its representation has minimum non-zero of "1"s. Therefore, its multiplication with an arbitrary input corresponds to utilization of minimum adders/substrators and shifters. Nevertheless, our proposed scheme with no adders/substrators, and by using only shifters in the multiplication block of the ASB-SVM implementation is able to perform faster than CSD multipliers.

3.2. FPGA implementation

To demonstrate the advantage and credibility of the proposed approach, we have implemented both MB-SVM and our designed ASB-SVM classifiers on Xilinx Virtex-6 LX240T FPGA and evaluated its hardware performance in terms of speed, accuracy and resource utilization. The dataset is constructed in MATLAB/SIMULINK environment. The training phase is done using MATLAB Classification Learner. Since our model is inherently a signal processing design, we have utilized MATLAB HDL Coder toolbox for efficient implementation. We have chosen the 16 bits fixed-point data format for value representation in the design. Fig. 3 illustrates one of the 15 blocks for the modulation classification with its corresponding input and output signals (according to $\frac{L(L-1)}{2}$). As can be seen from Fig. 3, the multiplications are replaced by arithmetic shift units that get two value: w which is the parameter value of the classifier and s that specifies how many shifts must be applied to w. The corresponding results are then accumulated through the adder tree. The classifier is pipelined in each stage to ensure the data flow credibility and high performance. The output of the adder tree summation is then compared with zero for further decision making in the voting scheme.

3.3. Results and Comparison

Table 1 shows the average accuracy of our proposed and the conventional approaches for signal classification which is verified in Simulink environment. As can be observed from Table 1, the overall accuracy of ASB-SVM is 98.2% in comparison to 98.9% from MB-SVM classifier which is a negligible loss. From Table 2, it is seen

Table 1. Average accuracy of the SB-SVM-based and the traditional SVM-based scheme for various SNR levels.

ASB-SVM	SNR					
Modulation	**0**	**5**	**10**	**15**	**20**	**25**
2-ASK	97.2	97.5	98.8	100	100	100
4-ASK	97.4	97.7	98.3	100	100	100
2-PSK	97.2	99.4	99.5	100	100	100
4-PSK	97.4	99.7	99.9	100	100	100
2-FSK	92.4	95.7	95.7	98.1	98.4	98.9
4-FSK	94.1	95.6	95.5	98.2	98.2	98.8
MB-SVM	SNR					
Modulation	**0**	**5**	**10**	**15**	**20**	**25**
2-ASK	98.2	99.2	99.4	99.8	100	100
4-ASK	98.1	99.1	99.6	100	100	100
2-PSK	98.2	99.3	99.3	99.5	100	100
4-PSK	98.1	98.4	98.6	98.9	100	100
2-FSK	95.3	97.2	98.7	98.1	100	100
4-FSK	95.4	96.1	97.4	98.1	99.4	99.8

Table 2. A comparison of resource utilization.

Resource	**ASB-SVM**	**MB-SVM**
Number of Slice Registers	9162 (3%)	7822 (2%)
Number of Slice LUTs	10392 (6%)	15262 (10%)
Number of Multiplications	0	165
Maximum frequency	384 MHz	200 MHz
Power consumption	2.92 W	3.66 W
Classify 200 example	0.5 ms	1 ms

that the number of LUT slice is reduced over 30% and the speed of the classification is increased by 45% without using multiplication. Moreover, our proposed ASB-SVM classifier is more accurate than the proposed system used in [18, 19], and also is capable of classifying a blind modulated signal within only 3 ns in the classifier block in a CRN which corresponds to over 300 million signals per seconds throughput and beats the HISTO-SVM proposed in [19] in terms of speed and overall accuracy in signal classification. Furthermore, the proposed HISTO-SVM in [19] is only applicable for modulated signals that has constellation characteristic which limits the utilization of such classification for other modulation schemes in CRNs. Moreover, comparing our designed system to deep learning-based shemes, the hardware implementation is more feasible specially in realtime applications and also it is not constrained to specific modulations.

4. CONCLUSION

We proposed the Arithmetic Shifter-Based SVM (ASB-SVM) architecture for realtime signal classification in a CRN, which is independent of the modulated signals. ASB-SVM benefits from replacing multiplications with shifters in its structure which enhances the speed of signal classification. Our designed scheme achieved over 30% resource reduction and 45% increase in speed in comparison to MB-SVM method. The classification of the modulated signals can be executed in parallel using multiple ASB-SVM classifiers. Such a design can boost the classification speed at the cost of more resource utilization. Furthermore, the proposed classifier is trained in a supervised manner. Our future direction is to promote a signal modulation classification with an unspervised learning-based classifier where the implementation can still benefits from shifters instead of multipliers.

5. REFERENCES

[1] M. Sh. Gupta and K. Kumar, "Progression on spectrum sensing for cognitive radio networks: A survey, classification, challenges and future research issues," *Journal of Network and Computer Applications*, vol. 143, pp. 47–76, 2019.

[2] A. Mohammadi, M. R. Taban, J. Abouei, and H. Torabi, "Fuzzy likelihood ratio test for cooperative spectrum sensing in cognitive radio," *ELSEVIER Signal Processing*, vol. 93, no. 5, pp. 1118–1125, 2013.

[3] Y. Wang, F. Karray, S. Kwong, K. N Plataniotis, H. Leung, M. Hou, E. Tunstel, I. J. Rudas, L. Trajkovic, O. Kaynak, , and J. Kacprzyk, "On the philosophical, cognitive and mathematical foundations of symbiotic autonomous systems (sas)," *arXiv preprint arXiv:2102.07617*, 2021.

[4] Y. Wang, M. Hou, K. N. Plataniotis, S. Kwong, H. Leung, E. Tunstel, I. J. Rudas, and L. Trajkovic, "Towards a theoretical framework of autonomous systems underpinned by intelligence and systems sciences," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 1, pp. 52–63, 2020.

[5] J. Liu, Sh. Jin, and W. Yue, "Performance evaluation and system optimization of green cognitive radio networks with a multiple-sleep mode," *Annals of Operations Research*, vol. 277, no. 2, pp. 371–391, 2019.

[6] M. Hasani-Baferani, J. Abouei, and Z. Zeinalpour-Yazdi, "Interference alignment in overlay cognitive radio femtocell networks," *IET Communications*, vol. 10, no. 11, pp. 1401–1410, 2016.

[7] G. Enemali, A. Adetomi, and T. Arslan, "A placement management circuit for efficient realtime hardware reuse on FPGAs targeting reliable autonomous systems," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2017, pp. 1–4.

[8] J. Fu, Ch. Zhao, B. Li, and X. Peng, "Deep learning based digital signal modulation recognition," in *The Proceedings of the Third International Conference on Communications, Signal Processing, and Systems*. Springer, 2015, pp. 955–964.

[9] T. Zare and J. Abouei, "Kernel-based generalized discriminant analysis for signal classification in cognitive radio," in *IEEE International Symposium on Telecommunications*. IEEE, 2014, pp. 1106–1112.

[10] T. O'shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[11] M. W. Aslam, Z. Zhu, and A. K. Nandi, "Automatic modulation classification using combination of genetic programming and knn," *IEEE Transactions on wireless communications*, vol. 11, no. 8, pp. 2742–2750, 2012.

[12] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4074–4077, 2019.

[13] S. Peng, H. Jiang, H. Wang, H. Alwageed, and Y. Yao, "Modulation classification using convolutional neural network based deep learning model," in *Wireless and Optical Communication Conference (WOCC)*. IEEE, 2017, pp. 1–5.

[14] G. J. Mendis, J. Wei, and A. Madanayake, "Deep learning-based automated modulation classification for cognitive radio," in *IEEE International Conference on Communication Systems (ICCS)*. IEEE, 2016, pp. 1–6.

[15] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 3, pp. 718–727, 2018.

[16] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.

[17] M. Barakat, W. Saad, and M. Shokair, "FPGA implementation of cyclostationary feature detector for cognitive radio ofdm signals," in *International Conference on Computer Engineering and Systems (ICCES)*. IEEE, 2018, pp. 215–218.

[18] M. A. Azza, A. El Moussati, and O. Moussaoui, "Implementation of an automatic modulation recognition system on a software defined radio platform," in *International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*. IEEE, 2018, pp. 1–4.

[19] C. Cardoso, A. R. Castro, and A. Klautau, "An efficient FPGA ip core for automatic modulation classification," *IEEE Embedded Systems Letters*, vol. 5, no. 3, pp. 42–45, 2013.

[20] S. Mosleh, J. Abouei, and M. Aghabozorgi, "Distributed opportunistic interference alignment using threshold-based beamforming in mimo overlay cognitive radio," *IEEE Trans. on Vehicular Technology*, vol. 63, no. 8, pp. 3783–3793, 2014.

[21] S. Fazeli-Dehkordy, J. Abouei, K. N. Plataniotis, and S. Pasupathy, "Markovian-based framework for cooperative channel selection in cognitive radio networks," *IET Communications*, vol. 8, no. 14, pp. 2458–2468, 2014.

[22] E. Azzouz and A. K. Nandi, "Automatic modulation recognition of communication signals," 2013.

[23] Y. Qi, Y. Wang, and C. Lai, "An improved SVM-based spatial spectrum sensing scheme via beamspace at low SNRs," *IEEE Access*, vol. 7, pp. 184759–184768, 2019.

A DRL BASED DISTRIBUTED FORMATION CONTROL SCHEME WITH STREAM-BASED COLLISION AVOIDANCE

Xinyou Qiu, Xiaoxiang Li, Jian Wang, Yu Wang, Yuan Shen

Beijing National Research Center for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
Email: {qxy18, lxx17}@mails.tsinghua.edu.cn, {jian-wang, yu-wang, shenyuan_ee}@tsinghua.edu.cn

ABSTRACT

Formation and collision avoidance abilities are essential for multi-agent systems. Conventional methods usually require a central controller and global information to achieve collaboration, which is impractical in an unknown environment. In this paper, we propose a deep reinforcement learning (DRL) based distributed formation control scheme for autonomous vehicles. A modified stream-based obstacle avoidance method is applied to smoothen the optimal trajectory, and onboard sensors such as Lidar and antenna arrays are used to obtain local relative distance and angle information. The proposed scheme obtains a scalable distributed control policy which jointly optimizes formation tracking error and average collision rate with local observations. Simulation results demonstrate that our method outperforms two other state-of-the-art algorithms on maintaining formation and collision avoidance.

Index Terms— Deep reinforcement learning, stream function, distributed control, collision avoidance.

1. INTRODUCTION

With the maturity and promotion of the Internet of Things, collaboration among agents becomes more and more imperative [1–3]. By deploying a multiple-agent system, sophisticated tasks such as harsh area exploration, disaster rescuing and map reconstruction become more tractable since each agent can be more concentrated on its own subtask. Among the key technologies of the multi-agent collaboration, formation control is the most important and practical one, through which we can assign each agent's relative position, (i.e. a desired formation "shape" [4]), and enable the multiple agents to explore the environment more efficiently. Furthermore, the robustness can be enhanced significantly by applying proper distributed operation and node failure tolerance methods.

In terms of formation control, collision avoidance is a critical task since agents often have to simultaneously maintain the desire formation and prevent collisions. Artificial potential field (APF) is a common solution for collision avoidance problems [5–7]. By constructing a potential field function, the agent can calculate the interaction forces with the obstacles and the destination. However, APF suffers from the local minima problem in complex environments with multiple obstacles, which could result in non-smooth motion or even be trapped in the saddle points [7]. Some researchers have focused on applying the concepts of fluid mechanics, such as stream function methods, to generate smoother trajectories in exclusion of local minimum [8, 9]. But the global position information of the agent and obstacles is required by these methods, which may be impractical in harsh cases. Deep reinforcement learning (DRL) [10–12] offers a promising solution to such a distributed formation control problem in arbitrary unknown systems due to its optimal-adaptive and model-free properties. By designing a proper cost function, agents can iteratively optimize policy by continuously exploring the environment even with local observation.

In this paper, we put forward a DRL scheme to train a formation control system with stream-based collision avoidance. The main contributions of our work are as follows:

1) We put forward a distributed formation control scheme with a modified stream-based collision avoidance policy. The proposed scheme requires only local observation for each agent and is easily achievable for real systems.

2) We design a DRL model, based on deep deterministic policy gradient (DDPG), to train the distributed policy in continuous action-state space. The decision-making and control layers are tightly coupled in the proposed model and thus guarantee a better collaboration between multiple agents. Simulation results show good performance compared with the existing obstacle-avoiding algorithms.

Notations: Throughout this paper, variables, vectors, and matrices are written as italic letters x, bold italic letters $\boldsymbol{x}$, and bold capital italic letters $\boldsymbol{X}$, respectively. Random variables and random vectors are written as sans serif letter x and bold letters $\mathbf{x}$, respectively. The notation $\mathbb{E}_{\mathbf{x}}\{\cdot\}$ is the expectation operator with respect to the random vector $\mathbf{x}$, and $\mathbb{1}(\cdot)$ is the indicator function which equals 1 if the condition is true and equals 0 otherwise.

2. PRELIMINARIES

In fluid dynamics, the stream function is often introduced to analyze a particle's behavior in a 2-D incompressible flow field. For example, streamlines plotted from different stream

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

functions represent the trajectories of particles inside the field. To obtain the stream function for an incompressible flow in a 2-D $x-y$ plane, we first consider the continuity equations given by

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0, \tag{1}$$

where u, v are the fluid velocity in x, y axis direction, respectively. The velocity field in a 2-D incompressible flows always satisfy (1). Define the function $\psi(x,y)$ as

$$u=\frac{\partial \psi}{\partial y}, v=-\frac{\partial \psi}{\partial x}, \tag{2}$$

then the continuity equation becomes

$$\frac{\partial}{\partial x}\frac{\partial \psi}{\partial y}+\frac{\partial}{\partial y}(-\frac{\partial \psi}{\partial x})=\frac{\partial^2 \psi}{\partial x \partial y}-\frac{\partial^2 \psi}{\partial y \partial x}=0. \tag{3}$$

Note that if there exists a function ψ which fulfills (3), it can be used to obtain streamlines $\psi = C$ for every point in the flow field. The points on the same streamlines share the same constant value C.

For a flow field with a cylinder-shaped obstacle locating at the origin, we can obtain the stream function by treating it as a combination of the uniform flow and the doublet flow [9]. Assume the strength of flow is U, the compound stream function is given as

$$\psi=\psi_{\text{uniform flow}}+\psi_{\text{doublet}}=Uy-U(\frac{r^2 y}{x^2+y^2}). \tag{4}$$

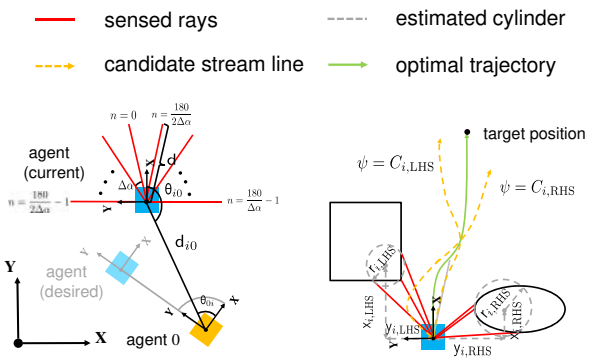
By following the assigned streamline, the agent can avoid collision smoothly since no local minima exists in the field.

3. SYSTEM MODEL

Consider a multi-agent navigation system with 1 virtual navigator agent and $N-1$ follower agents moving in a 2-D Euclidean plane at time $k \in [0, T_{\max})$ with the following dynamic model

$$\begin{aligned}\mathbf{x}_i(k+1) =&[x_i(k)+\Delta x_i(k), y_i(k)+\Delta y_i(k), v_i(k)+ka_i(k),\\ &\alpha_i(k)+k\omega_i(k), \omega_i(k)+k\beta_i(k)]^\intercal+\mathbf{w}_i(k)\\ \triangleq& f(\mathbf{x}_i(k),\boldsymbol{u}_i(k))+\mathbf{w}_i(k)\end{aligned} \tag{5}$$

where $i=0,...,N-1$. $i=0$ is the virtual navigator agent and the rest are follower agents. $\boldsymbol{p}_i(k)=[x_i(k),y_i(k)]^\intercal \in \mathcal{R}^2$ and $\alpha_i \in (-\pi,\pi]$ are the position and the orientation of agent i in the global coordinate system, $v_i(k)$ and $\omega_i(k)$ are the velocity and the angular velocity. $\Delta x_i(k)=\frac{v_i(k)}{\omega_i(k)}[\sin(\alpha_i(k)+k\omega_i(k))-\sin(\alpha_i(k))]$, $\Delta y_i(k)=\frac{v_i(k)}{\omega_i(k)}$ $[\cos(\alpha_i(k)+k\omega_i(k))-\cos(\alpha_i(k))]$, $\boldsymbol{u}_i(k)=[a_i(k),\beta_i(k)]^\intercal \in \mathcal{R}^2$ are the control inputs of the acceleration and angular acceleration, and $\mathbf{w}_i(k) \sim \mathcal{N}(0,\boldsymbol{C})$ are independent white Gaussian state noises with zero means and covariance matrix $\boldsymbol{C}$ in the local coordinate system. System (5) is assumed stabilizable, i.e. there exists a continuous control $\boldsymbol{u}_i$ such that the system is asymptotically stable. The communication graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$ is assumed as an undirected



(a) system model and geometric relationship (b) proposed stream-based avoiding policy

Fig. 1. The schematic diagram of our system and method.

connected graph if the adjacent agents stay in the connection zone δ_i. The set of neighbors of $i \in \mathcal{V}$ is defined as $\mathcal{N}_i := \{j \in \mathcal{V} : (i,j) \in \mathcal{E}\}$. For any agent i, the relative position set is denoted as $\mathcal{P}_i=\{[d_{ij},\theta_{ij}]^\intercal, \forall j \in \mathcal{N}_i\}$, which includes the distance and angle to all adjacent agents that can be communicated through antenna array-based sensors. Agent 0 will continuously broadcast $\theta_{0j}(k)$ to every connectable agent through communication. To detect the collision, lidar-based distance sensors are equipped on the follower agents. The sensors provide distance measurements $\mathbf{d}=\boldsymbol{d}+\mathbf{n}$ with an angle resolution of $\Delta\alpha$, where $d_{\min} \le d \le d_{\max} \forall d \in \boldsymbol{d}$ and $\mathbf{n} \sim \mathcal{N}(0,\boldsymbol{C}_{\text{lidar}})$ denotes additive white Gaussian noises with zero means and covariance matrix $\boldsymbol{C}_{\text{lidar}}$, as Fig. 1(a) illustrates.

Agent 0 is responsible for providing a trajectory which navigates the swarm toward the goal. The follower agents must maintain their predefined relative position $\boldsymbol{\eta}_i=[d_i\cos(\theta_i), d_i\sin(\theta_i)]^\intercal$ with respect to agent 0 while avoiding collision. Define the tracking error for agent i as

$$\boldsymbol{e}_i(k)=\boldsymbol{p}_i(k)-\boldsymbol{p}_0(k)-\boldsymbol{\eta}_i. \tag{6}$$

Note that $d_{i0}(k)$ and $\theta_{i0}(k)$ are available through agent 0's broadcast, (6) can be rewritten as following form

$$\boldsymbol{e}_i(k)=\boldsymbol{z}_i(k)-\boldsymbol{\eta}_i \tag{7}$$

where

$$\begin{aligned}\boldsymbol{z}_i(k)&=\boldsymbol{p}_i(k)-\boldsymbol{p}_0(k)\\ &=[d_{0i}(k)\cos(\theta_{0i}(k)), d_{0i}(k)\sin(\theta_{0i}(k))]^\intercal\end{aligned} \tag{8}$$

is the relative displacement vector between follower agent i and agent 0. Although $\boldsymbol{p}_0(k), k=0,...,T_{\max}$ is unknown for agent i, tracking the trajectory is viable through the observed relative position. To minimize the tracking error, we introduce the following cost function [13]

$$r_{i,\text{tracking}}(k)=\boldsymbol{e}_i^\intercal(k)\boldsymbol{Q}_e\boldsymbol{e}_i(k), \tag{9}$$

where $\boldsymbol{Q}_e$ is a positive definite matrix.

4. APPROACH AND IMPLEMENT DETAILS

4.1. Distributed stream-based collision avoidance

Traditionally, stream-based collision avoidance requires the agent to follow a set of virtual leader's trajectory on the de-

sired streamline. This set will be set at the beginning according to the coordinates of the obstacle. However, most unmanned vehicles are designed to explore unknown areas where prior knowledge of the obstacles is unavailable. Therefore, we introduce a safe distance range $\boldsymbol{d}_{\text{safe}} = [d_{\text{risk}}, d_{\text{stop}}]^\intercal$, where d_{risk} is the risk distance that the agent should start the avoiding behavior, and d_{stop} is the minimum braking distance for an agent to stop from the highest speed [14]. To apply stream-based collision avoidance on our system, several assumptions are necessary:
1) The velocity of the agent is nonnegative, i.e., $v_i(k) \geq 0$,
2) The minimum horizontal projection length of obstacles is at least $2d_{\text{risk}}\sin(\Delta\alpha)$, so that it can be detected by at least three rays.

When an obstacle is detected by agent i, the direction of detected rays form a discrete interval whose indices can be denoted as a set $\mathcal{A}_i = \{n_{\text{start}}, n_{\text{start}}+1, ..., n_{\text{end}}-1, n_{\text{end}}\}$, where $|\alpha_n - \alpha_{n-1}| = \Delta\alpha, \forall n > 0$. The endpoint of each ray n is denoted as $\mathbf{p}_n(k) = [\mathsf{d}_n(k)\cos(\alpha_n(k)), \mathsf{d}_n(k)\sin(\alpha_n(k))]^\intercal, n \in [n_{\text{start}}, n_{\text{end}}]$, where $n_{\text{start}}, n_{\text{end}}$ are the indices of the start and end ray in the interval. We can always find a shortest ray $m = \arg\min ||\mathbf{p}_m(k)||, m \in \mathcal{A}_i \backslash \{n_{\text{start}}, n_{\text{end}}\}$ in the interval, where m represents the index of the shortest ray except the start and end ray. The three endpoints $\mathbf{p}_m(k)$, $\mathbf{p}_{n_{\text{start}}}(k)$ and $\mathbf{p}_{n_{\text{end}}}(k)$ can form a triangle as long as they are not on a line.

To deal with multiple obstacles, we divide the front semicircle into left-hand side (LHS) and right-hand side (RHS) to represent two flow fields, whose interval sets are denoted as $\mathcal{A}_{i,\text{LHS}} = \{n \in \mathcal{A}_i | \alpha_n > 0\}$ and $\mathcal{A}_{i,\text{RHS}} = \{n \in \mathcal{A}_i | \alpha_n < 0\}$. The two fields are independent and consider only the foremost interval sets (minimum n_{start} on either side). By finding the circumcenter of the triangle, we can obtain a virtual cylinder with radius $\mathbf{r}_{i,\text{obstacle}} = [\mathsf{r}_{i,\text{LHS}}, \mathsf{r}_{i,\text{RHS}}]^\intercal$ and relative position $\mathbf{p}_{\text{i,cyl}}(k) = [\mathbf{x}^\intercal_{\text{i,cyl}}(k), \mathbf{y}^\intercal_{\text{i,cyl}}(k)]^\intercal$ where $(\cdot)_{i,\text{cyl}}(k) = [(\cdot)_{i,\text{LHS}}, (\cdot)_{i,\text{RHS}}]^\intercal$ under the agent's coordinate system, as Fig. 1(b) illustrates.

Based on the assumptions, it is impossible for any obstacle to be ahead of the agent after being avoided, i.e. during avoidance, the virtual cylinder is pretended identical if $|(\alpha_{n_{\text{start}}}(k))| > |(\alpha_{n_{\text{start}}}(k-1))|$. Given that the cylinders being the same, $-\mathbf{p}_{i,\text{cyl}}(k)$ can represent the displacements of agent i to the origin of virtual cylinders with radius $\mathbf{r}_{i,\text{obstacle}}(k)$. To this end, the agent is capable of calculating $\mathbf{c}_i = [\mathsf{c}_{i,\text{LHS}}(k), \mathsf{c}_{i,\text{RHS}}(k)]^\intercal$ in (4) with local observation.

The desired stream values of the two flow fields are denoted as $\boldsymbol{c}_{i,\text{desired}} = [C_{i,\text{LHS}}, C_{i,\text{RHS}}]^\intercal$. To ensure the availability of the streamlines for different agent sizes, $\boldsymbol{c}_{i,\text{bound}}$ is given based on $[x_{i,\text{LHS}}, y_{i,\text{LHS}}, x_{i,\text{RHS}}, y_{i,\text{RHS}}]^\intercal = [0, -d_{\text{stop}}, 0, d_{\text{stop}}]^\intercal$. From [8], we have observed that following the streamline is in fact to find $\mathbf{c}_i = \min ||\boldsymbol{c} - \boldsymbol{c}_{i,\text{desired}}||$. On top of that, our algorithm tracks $\boldsymbol{c}_{i,\text{desired}}$ directly instead of predefined point sets of the streamline. The avoidance cost function can be formulated as

$$r_{i,\text{avoiding}}(k) = \sum_{s=0}^{\text{len}(\boldsymbol{c}_i)-1} \mathbb{1}(avoid_s(k))(\mathbf{c}_i(s) - \boldsymbol{c}_{i,\text{desired}}(s))^2 \left(\frac{1}{||\mathbf{p}_m(k)||} - \frac{1}{d_{\text{risk}}}\right), \tag{10}$$

where $avoid_s(k)$ is the avoidance flag, $(\mathbf{c}_i(s) - \boldsymbol{c}_{i,\text{desired}}(s))^2$ calculates the stream value error. The cost is multiplied with a potential field-based cost [5] in case $\mathbf{c}_i(s)$ changes too subtly when $\mathbf{y}_{i,\text{cyl}}(s) \approx 0$. (10) reliefs the strong penalty when the agent tries to go through a faster path but with more obstacles. Thus, the agent can recover the formation faster while remaining safe in a multi-obstacle environment. The entire policy is shown in Algorithm 1.

Algorithm 1 Avoidance Decision Policy

```
obtain distance measurements of agent i
for s in 0, ..., len(c_i) − 1 do
    initialize avoid_s(k) as false
    set avoid_s(k) true and calculate p_{i,cyl}(s), r_{i,obstacle}(s) if detected
    if avoid_s(k) then
        if not avoid_s(k − 1) then
            calculate c_{i,desired}(s) by (4), check availability with c_{s,bound}
        else
            calculate c_i(s)
            update c_{i,desired}(s) ← c_i(s) if new obstacle detected
        end if
    end if
end for
return r_{i,avoiding}(k) in (10)
```

Remark 1 *To apply the algorithm, default values should be set if no triangle can be formed by the detected results. Also the agent should choose a side to avoid once the obstacle overlaps LHS and RHS.*

4.2. DRL-based control policy optimization

We apply DDPG [15] to optimize the control policy without knowing the system model. A critic network $Q^{\boldsymbol{\varpi}}(\mathbf{o}, \boldsymbol{u})$ and an actor network $\boldsymbol{u}^{\boldsymbol{\phi}}(\mathbf{o})$, parameterized by $\boldsymbol{\varpi}, \boldsymbol{\phi}$, are created to evaluate the value of the current state and produce the control inputs.

The critic network of the entire system $Q(k)$ is defined as

$$Q(k) = \sum_{i=1}^{N-1} Q_i^{\boldsymbol{\varpi}}(\mathbf{o}_i(k), \boldsymbol{u}_i^{\boldsymbol{\phi}}(k)) = \sum_{i=1}^{N-1} \sum_{t=k}^{T_{max}} \gamma^{t-k} r_i(\mathbf{o}_i(k)), \tag{11}$$

where γ is the discount factor of expected future cost, $r_i(\mathbf{o}_i(k)) = r_{i,\text{tracking}}(k) + r_{i,\text{avoiding}}(k)$ is the sum of (9) and (10), $\mathbf{o}_i(k) = [\boldsymbol{e}_i(k), \alpha_i(k), v_i(k), \omega_i(k), \mathcal{P}_i, \mathcal{P}_{\text{sensor}}, \mathbf{o}_{\text{cyl}}]^\intercal$ is the observation vector, $\mathcal{P}_{\text{sensor}} = \{[\mathsf{d}_n(k), \alpha_n(k)]^\intercal, \forall n = \{n_{\text{start}}, m, n_{\text{end}}\} \in \{\mathcal{A}_{i,\text{LHS}}, \mathcal{A}_{i,\text{RHS}}\}\}$ includes the chosen detected results from the two sides. To promote a faster convergence, $\mathbf{o}_{\text{cyl}}(k) = [\mathbf{p}_{\text{cyl}}(k), \mathbf{r}_{\text{obstacle}}(k)]^\intercal$ is preprocessed

Table 1. Simulation settings

	name	value
environment settings	agent shape(size)	square(0.1m)
	obstacle size	$r_{\text{obstacle}} \sim \mathcal{U}(0.1, 0.5)$m
	lidar resolution	$\Delta\alpha = 3^\circ$
	detect range, noise	$\boldsymbol{d} \in [0, 2\text{m}], \mathbf{n} \sim \mathcal{N}(0, (0.2\text{m})^2\boldsymbol{I})$
	maximum $v_i(v_0)$	0.5m/s(0.35m/s)
	maximum $\lvert\omega_i\rvert(\lvert\omega_0\rvert)$	$\lvert$0.2rad/s$\rvert(\lvert$0.06rad/s$\rvert)$
	connection zone	$\delta_i = 7$m
	safe distance $\boldsymbol{d}_{\text{safe}}$	$[d_{\text{risk}} = 0.7\text{m}, d_{\text{stop}} = 0.4\text{m}]^\intercal$
hyperparameters	network architecture	$(64, 128, 128)$ FC layers
	critic (actor) lr	$10^{-3}(10^{-4})$
	action space	$u^\phi = [u_0, u_1, u_2]^\intercal$ with softmax layers
	batch size	1024
	training episodes	30000
	episode steps	250
	step time	0.1sec

Table 2. Evaluation results

Method	Tracking error (m)	Collision rate $(\%)$
APF, d_{risk}	2.7704	6.46
APF, d_{stop}	1.9538	5.56
Our method	**1.3358**	**0.93**

from $\mathcal{P}_{\text{sensor}}$ instead of learning it by DDPG. Since the formation system is homogeneous [12] and we deploy the same policy on all follower agents, the policy can be deployed distributedly. The minimization of (11) and the policy is obtained by updating the network with [15].

5. SIMULATION RESULTS

The proposed method is evaluated in an obstructive area. At most 5 round obstacles are randomly scattered in a $14 \times 14\text{m}^2$ area around agent 0, who provides the trajectory with navigation method in [16]. 4 follower agents are deployed around agent 0 to form a circle with radius 2.1m. The state transfer noise is $\mathbf{w} \sim \mathcal{N}(0, \boldsymbol{C})$, where $\boldsymbol{C} = \text{diag}\{10^{-2}\text{m}^2, 10^{-2}\text{m}^2, 10^{-4}(\text{m/s})^2, 3.2 \times 10^{-4}\text{rad}^2, 3.2 \times 10^{-6}(\text{rad/s})^2\}$. The rest environment settings and the hyperparameters of DDPG are shown in Table. 1. The final control input is $\boldsymbol{u} \triangleq [a = u_0, \beta = (u_1 - u_2)]^\intercal$, which is mapped to proper value in case the speed limit is exceeded.

Our method is compared with APF [5] and the exponential reward proposed by Chao Wang, et al [16]. The convergence curves of different methods are shown in Fig. 2. Since Chao Wang's reward is designed for large-scale obstacles, it fails to avoid scattered obstacles in our scene. APF in different safe distances and Stream-based method both ensure collision-free for more than 98% during training, yet the later converges faster and achieves lower tracking error.

To evaluate the robustness of our method, we increase the episode length to 2500 steps and run 100 times per method to obtain an average result. Table. 2 shows the average tracking error and collision rates of three models during evaluation. Fig. 3 depicts the entire tracking errors under the same obstacles deployment, in which the fluctuations are owing to

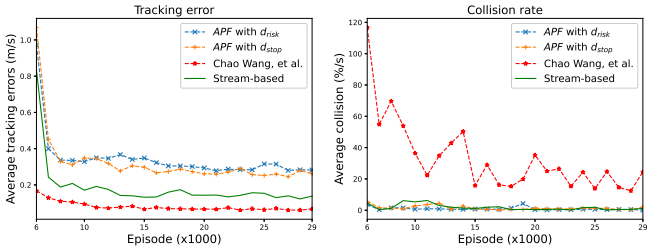


Fig. 2. Training curves of formation tracking with different avoidance cost, ignoring the warmup stages.

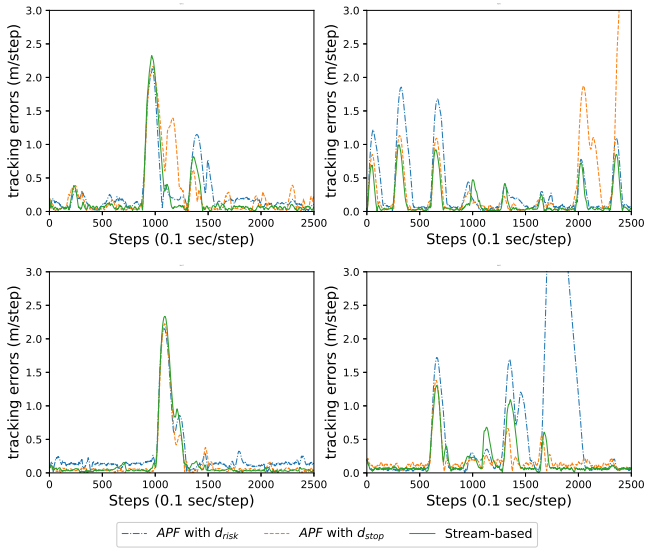


Fig. 3. The tracking errors under fixed environment.

avoidance. Through the elimination of unnecessary avoiding penalties, our method can return the correct formation position more efficiently without divergence. Thus, the robustness of our model when encountering multiple obstacles under long task duration is guaranteed. The entire simulation trajectories of the three models are shown in `https://youtu.be/jpsQ-kBJZk8`.

6. CONCLUSION AND FUTURE WORK

In this paper, we propose a collision avoidance scheme for a formation navigation system based on the stream function. Unlike traditional stream-based methods, the proposed scheme avoids the requirement of global information by estimating the virtual flow field based on agents' local sensors. The numerical result reveals the improvement of our scheme in both collision rate and tracking error for 5% and 0.6m, respectively. In future work, we will investigate the potential of cooperation among the agents, along with the deployment of our algorithm on the real robot swarm system.

Acknowledgment

This research was supported by the National Natural Science Foundation of China under Grant 61871256 Tsinghua University Initiative Scientific Research Program National Key R&D Program of China 2020YFC1511803

7. REFERENCES

[1] Yifeng Xiong, Nan Wu, Yuan Shen, and Moe Z. Win, "Efficiency of cooperation and its geometric interpretation in network localization," in *Proc. IEEE Int. Conf. Commun. Workshop*, Shanghai, China, May 2019, pp. 1–6.

[2] Moe Z. Win, Andrea Conti, Santiago Mazuelas, Yuan Shen, Wesley M. Gifford, Davide Dardari, and Marco Chiani, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.

[3] Amir Amini, Amir Asif, and Arash Mohammadi, "CEASE: A Collaborative Event-Triggered Average-Consensus Sampled-Data Framework With Performance Guarantees for Multi-Agent Systems," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6096–6109, Dec 2018.

[4] Yuanpeng Liu, Yunlong Wang, Jian Wang, and Yuan Shen, "Distributed 3D relative localization of UAVs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11756–11770, Oct. 2020.

[5] Carlos Sampedro, Hriday Bavle, Alejandro Rodriguez-Ramos, Paloma de la Puente, and Pascual Campoy, "Laser-Based Reactive Navigation for Multirotor Aerial Robots using Deep Reinforcement Learning," in *2018 Int. Conf. Intell. Robots and Syst.*, 2018.

[6] Bohao Li and Yunjie Wu, "Path Planning for UAV Ground Target Tracking via Deep Reinforcement Learning," *IEEE Access.*, vol. 8, pp. 29064–29074, Feb 2020.

[7] Makiko Okamoto and Maruthi R. Akella, "Novel potential-function-based control scheme for non-holonomic multi-agent systems to prevent the local minimum problem," *Int. J. Syst. Sci.*, vol. 46, no. 12, pp. 2150–2164, Nov 2015.

[8] Stephen Waydo and Richard M. Murry, "Vehicle motion planning using stream functions," in *IEEE Int. Conf. Robotics and Automation.*, 2003, pp. 2484–2491.

[9] QiangWang, Jie Chen, Hao Fang, and Qian Ma, "Flocking control for multi-agent systems with stream-based obstacle avoidance," *Trans. Institute of Measurement and Control.*, vol. 36, pp. 391398, 2014.

[10] Said G. Khan, Guido Herrmann, Frank L. Lewis, Tony Pipe, and Chris Melhuish, "Reinforcement learning and optimal adaptive control: An overview and implementation examples," *Annual Reviews in Control.*, vol. 36, no. 1, pp. 42–59, Jun 2012.

[11] Richard S. Sutton, Andrew G. Barto, and Ronald J. Williams, "Reinforcement learning is direct adaptive optimal control," *IEEE Control Syst. Mag.*, vol. 12, no. 2, pp. 19–22, Apr 1992.

[12] Guoxing Wen, C. L. Philip Chen, Jun Feng, and Ning Zhou, "Optimized Multi-Agent Formation Control Based on an IdentifierActorCritic Reinforcement Learning Algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 2719–2731, Oct 2018.

[13] Hao Liu, Qingyao Meng, Fachun Peng, and Frank L. Lewis, "Heterogeneous formation control of multiple UAVs with limited-input leader via reinforcement learning," *Neurocomputing.*, vol. 412, pp. 63–71, June 2020.

[14] Jong Hun Park and Uk Youl Huh, "Path Planning for Autonomous Mobile Robot Based on Safe Space," *IEEE J. Electr. Eng. Technol.*, vol. 11, no. 5, pp. 1441–1448, Mar 2016.

[15] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra, "Continuous Control with Deep Reinforcement Learning," in *Int. Conf. Learning Representations.*, 2016.

[16] Chao Wang, Jian Wang, Yuan Shen, and Xudong Zhang, "Autonomous Navigation of UAVs in Large-Scale Complex Environments: A Deep Reinforcement Learning Approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2124–2136, Mar 2019.

MATCHING MODELS FOR CROWD-SHIPPING CONSIDERING SHIPPER'S ACCEPTANCE UNCERTAINTY

Shixuan Hou, and Chun Wang, Member, IEEE,

Concordia University

ABSTRACT

Crowd-shipping systems, which use occasional drivers to deliver parcels with compensations, offer greater flexibility and cost-effectiveness than the conventional company-owned vehicle shipping system. This paper investigates a dynamic crowd-shipping system that uses in-store customers as crowd-shippers to deliver online orders on their way home under the condition that the crowd-shippers' acceptances are uncertain. Optimal matching results between online orders and crowd-shippers and optimal compensation schemes should be determined to minimize the total costs of the crowd-shipping system. To this aim, we formulate this problem as a two-stage optimization model that determines matching results and compensation schemes sequentially. To evaluate the proposed optimization model, we conduct a series of computational experiments. Results show that the average delivery cost is reduced by 7.30 %, compared to the conventional shipping system.

Index Terms— Crowd-shipping, DES, in-store customer, online order, cost minimization

1. INTRODUCTION

The spectacular growth of e-commerce promotes many companies to seek a creative and innovative solution to provide fast, cheap, and reliable delivery services to final consumers [1]. "Crowd-shipping" is emerged from this opportunity. It is an innovative delivery strategy that uses ordinary people to perform same-day delivery rather than professional drivers or delivery companies (e.g., UPS or FedEx). Some big companies like Wal-Mart [2], DHL [3], and AmazonFlex [4] discussed this new concept. Specifically, Wal-Mart uses in-store customers to deliver parcels on their way home, DHL utilizes extra capacities of residents along their daily routes, while AmazonFlex drivers have to pick up parcels from stations and deliver them to final consumers. All the companies encourage ordinary people to "carpool" a parcel that they pick parcels up and deliver the goods on the way to their original destination with being paid a small compensation to reimburse their extra travel costs[1]. Due to the utilization of extra capacities of ordinary people, crowd-shipping is a social-economical friendly strategy that dramatically reduces urban traffic congestion and carbon emission and, at the same time, reduces the delivery costs of retailers or logistics companies. Over the past few years, considering the crowd-shipping strengths, many online platforms, like Jing Dong [1], and PiggyBee[2], have been dedicated to implementing it.

In this paper, we focus on a crowd-shipping system that uses in-store customers (the crowd-shippers) to supplement professional drivers to deliver online orders on the way to their destinations. Moreover, we consider a matching problem formulated as a single-order single-shipper assignment that allows crow-shippers to reject assigned delivery tasks. Methodologically, this paper makes three significant contributions. The first contribution is that we propose a two-stage stochastic optimization model. The first stage model determines optimal matching results between online orders and crowd-shippers that satisfy as many online orders as possible. And the second stage model is to choose an optimal compensation scheme that minimizes total delivery costs. The second contribution is that each crowd-shipper is free to accept or reject assigned online orders, so we propose a DES simulation framework introducing two events representing each crowd-shipper's acceptance and rejection behaviors.

Furthermore, we also consider the dynamic characteristics that crowd-shippers may enter and leave the store at any time; therefore, we introduce two events representing the DES framework's mentioned behaviors. To cope with the dynamic matching problem in the crowd-shipping system, we repeatedly solve the proposed two-stage optimization problem each time a new crowd-shipper check out. The third contribution is that we consider the influences of distance and compensation on each crowd-shipper's decision, formulated as a discrete choice model.

We organize the remainder of the paper as follows. In Sec. 2, we survey the state of the art. And the dynamics of the considered crowd-shipping system and the mathematical formulation of the optimal matching problem are provided in Sec. 3 and Sec. 4; We present the results of computational experiments in Sec. 5; In Sec. 6, we summarize the significant contributions and give the possible extension of our work.

[1] https://corporate.jd.com/

[2] https://www.piggybee.com/en/

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

2. LITERATURE REVIEW

This paper focuses on a crowd-shipping system that uses in-store customers to deliver online orders. The major challenge of the system is determining appropriate matching results that can reduce total delivery costs. In this framework, Archetti *et al.* [5] propose a multi-start heuristic approach to match in-store customers and online orders to minimize total delivery costs. Gdowska *et al.* [1] introduce a heuristic algorithm to increase the subset of orders to be proposed to occasional couriers up to the point where total gains are the (locally) highest. And Kafle *et al.* [6] uses pedestrians or cyclists' capacities to perform last-leg parcel delivery; this paper proposes a mixed-integer programming model and uses a Tabu search algorithm to determine the matchings between crowd-shippers and parcels. Wang *et al.* [7] proposes various pruning techniques to help solve a large-scale citizen workers-delivery tasks assignment problem, which is formulated as a network min-cost flow problem, by reducing the size of network size. Moreover, the proposed methods' efficiency is verified by conducting comprehensive experiments with accurate data of Singapore and Beijing. Wang *et al.* [8] also proposes a heuristic algorithm to assign delivery tasks to submitted car trips, aiming to maximize the total utility of the crowd-shipping system.

In conclusion, the mentioned research investigates the static version of crowd-shipping systems; however, crowd-shipping systems are dramatically changing over time in the real world. To cope with the gap, Dayarian *et al.* [9] propose two rolling horizon approaches to match in-store customers with online orders repeatedly; their study contributes to the second rolling horizon approach that they consider the uncertainties about the number of future orders and customers. Similarly, Arslan *et al.* [10] consider a dynamic environment where both new delivery tasks and driver trips arrive over time paper formulates the matching problem as a mixed-integer programming model and proposes a recursive heuristic algorithm to solve it. And a dynamic matching algorithm, proposed by [11] assigns parcels into submitted travel plans based on the transportation routes and time constraints. Also, Chen *et al.* [12] proposes an idea that uses taxi drivers to perform the last-mile delivery in a real-time decision-making environment.

Another significant challenge of crowd-shipping is determining an appropriate compensation scheme to balance the number of engaged crowd-shippers and total delivery costs. However, the limited number of researches focus on the issue; most of them assume that their crowd-shipping systems adopt one of the following four pricing strategies (as shown in Table 1). To the best of our knowledge, only Archetti *et al.* [5] compares the impact of different compensation schemes on the cost-effectiveness of the crowd-shipping system. However, this paper does not involve a more in-depth exploration of the pricing strategy. Furthermore, all of the collected literature in this work assume crowd-shippers will accept all delivery tasks assigned to them. Except for Gdowska *et al.* [1], they point out that in-store customer is free to accept or reject the assigned tasks; however, this paper does not profoundly investigate the impact of both distance and compensation on the probability of acceptance of an assigned delivery task. To cover these gaps, we consider an in-store customer-based crowd-shipping system. And we propose a two-stage stochastic optimization model to determine the matchings between online orders and in-store customers and an optimal compensation scheme to minimize the total delivery costs considering the uncertainty of in-store customers' choices. Besides, to estimate the performances of the proposed crowd-shipping system in this paper, a Discrete Event System (DES) simulation model is built in Sec 3.

Pricing strategy	Reference
Fixed fee + extra time cost	[9][11]
Fixed fee + extra distance cost	[7] [8]
Fixed fee per item	[5] [10] [1]
Fixed fee per hour	[12] [6]

Table 1. Pricing strategy classification

3. MODELING

This paper considers a crowd-shipping system that uses in-store customers, as crowd-shippers, to deliver online orders for a store. Each crowd-shipper may enter the store, go shopping, leave the store with or without taking an online order. Moreover, each crowd-shipper must declare his or her trip destination, which benefits the system manager to assign appropriate online order to the crowd-shipper. Since the dynamic characteristics of each crowd-shippers, we model the considered crowd-shipping system as a Discrete Event System. We organize this section as follows: after describing the crowd-shipping DES model, some significant events characterizing the system dynamics are described in detail in Sec 3.2 and Sec. 3.3.

3.1. DES modeling of crowd-shipping

The events affecting the states of crowd-shippers are e_i^v, e_i^p, e_i^a, e_i^n, e_i^d, and e_i^r which represent the crowd-shipper arrival, crowd-shipper check-out, assigned online order acceptance, rejection, online order delivery, and destination arrival. Moreover, the state space Y consists of the set of possible values of the vector $\boldsymbol{y}(k) = [y_1(k)\ y_2(k)\ ...\ y_{|I|}(k)]^T$, where the generic entry $y_i(k)$ represents the state of the crowd-shipper i, after the occurrence of the k^{th} event. The value of the state variable $y_i(k)$ indicates if the crowd-shipper i:

- is not in the store: $y_i(k) = y_0$;
- is in the store: $y_i(k) = y_1$;

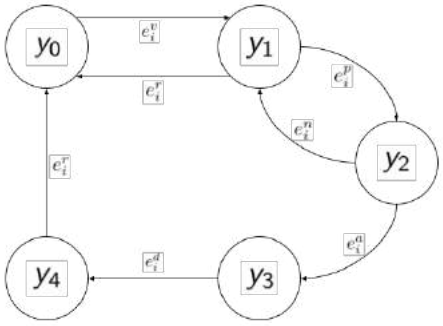


Fig. 1. State transition diagram of each in-store customer

- is evaluating an assigned online order: $y_i(k) = y_2$;
- has already accepted an online order: $y_i(k) = y_3$;
- has already delivered an online order: $y_i(k) = y_4$.

The state transition diagram is showed in Fig. 1. Nevertheless, it should be noted that in the model, we assume that when a crowd-shipper checks out, the manager of the system propose an assignment result to the crowd-shipper, after the crowd-shipper makes the final decision, the crowd-shipper must leave the store immediately.

3.2. Crowd-shipper arrival event

The occurrence of crowd-shipper arrival event represents the change of crowd-shipper state from y_0 to y_1. As regards the event timing, the arrival event e_i^v is assumed to occur at τ_i^v; regarding the system state, the occurrence of the arrival event of the crowd-shipper i implies the scheduling of the relevant check-out event e_i^p at $\tau_i^p = \tau_i^v + t$, being $t \sim \mathcal{N}(\mu, \sigma^2)$, a Gaussian stochastic variable of which expectation μ equal to the average shopping time in a store and variance is equal to σ^2. Moreover, we assume that crowd-shipper arrivals follow Poisson Process; the elapsed time between two arrival events results to be an exponential stochastic variable with the rate λ_1 crowd-shippers per hour.

3.3. Acceptance and rejection of assigned tasks

The crowd-shipper check-out event e_i^p, occurring at τ_i^p, triggers the solution of the first-stage optimization problem described in Sec. 4, based on the solution, the system proposes delivery tasks to crowd-shippers with compensations in the form of cash. The crowd-shipper is free to choose whether to accept the proposal or reject it. In the considered model, this freedom is defined as the acceptance probability p_i^a, determined by the solution of the second-stage optimization problem described in Sec. 4. Therefore, the event e_i^a is scheduled with probability p_i^a and the event e_i^n is scheduled with probability $p_i^n = 1 - p_i^a$.

4. OPTIMIZATION MODEL

In the section, the mathematical programming problem triggered by the event e_i^p is defined. In this framework, the proposed problem determines, on one side, the optimal assignments between crowd-shipper and delivery task for satisfying as many online orders as possible, and on the other side, the optimal compensation provided to crowd-shipper who is willing to deliver parcels and maximize the total saving of delivery costs of the store. The proposed optimization problem can be formulated as a constrained non-linear mixed-integer programming problem. Since the presence of the integer variables, an analytical solution approach can not be applied.

To cope with this optimization problem, we proposed a heuristic solution that decomposes the whole assignment problem with compensation into two sub-problems: the first one determines the optimal assignment between online orders and crowd-shippers by considering the online order satisfaction rate; the second one maximizes the total saving of the whole system by determining an optimal compensation scheme. Note that the solution of the first sub-problem is the input of the second one. And it is worth mentioning that because of the decomposition of the whole problem, the optimality of the solution can not be guaranteed, but it results in high computational efficiency and can be applied in a dynamic market environment.

The formulation of the two stages is presented in detail in Sec. 4.1 and Sec. 4.2, respectively. And the notation of the optimization model is shown in Table 4.

4.1. First stage optimization

In the first stage optimization problem, appropriate online orders are assigned to appropriate crowd-shippers to satisfy as many online orders as possible and reduce the total extra travel distance of all crowd-shippers. Below, the optimal assignment problem is formulated as a linear binary programming problem:

$$\min_{\boldsymbol{x}} \omega_1 G_1(\boldsymbol{x}) + \omega_2 G_2(\boldsymbol{x}) \tag{1}$$

$$G_1(\boldsymbol{x}) = \frac{|J| - \sum_{\forall i \in I} \sum_{\forall j \in J} x_{i,j}}{|J|} \tag{2}$$

$$G_2(\boldsymbol{x}) = \sum_{\forall i \in I} \sum_{\forall j \in J} x_{i,j} \frac{D_{i,j}}{D(o, d_i)} \tag{3}$$

subject to

$$D_{i,j} = \mathcal{D}(o, d_j) + \mathcal{D}(d_j, d_i) - \mathcal{D}(o, d_i) \tag{4}$$

$$\sum_{\forall j \in J} x_{i,j} \leq 1, \ \ \forall i \in I \tag{5}$$

$$\sum_{\forall i \in I} x_{i,j} \leq 1, \ \ \forall j \in J \tag{6}$$

Notation	Meaning
i	Indicator of a generic crowd-shipper $\in I$
I	Set of all crowd-shippers
$\hat{I}$	Set of matched crowd-shippers
j	Indicator of a generic online order $\in J$
J	Set of all online order
$\hat{J}$	Set of matched online order
o	Store address
$\mathcal{D}(a,b)$	Distance between zone a and zone b
$x_{i,j}$	Binary variable that indicates the matching between i and j
$d_{i(j)}$	Destination of shipper i (order j)
$D_{i,j}$	Extra distance i has to travel to deliver j
$U_i^{a(n)}$	Utility of shipper i to accept/reject to deliver online order
S	Set of all compensations
s_i	The compensation paid to shipper i
c_j	Original delivery cost of online order j

Table 2. Considered notation in the proposed optimization framework of crowd-shipping system

$$x_{i,j} = \{0,1\}, \quad \begin{cases} \forall i \in I \\ \forall j \in J \end{cases} \tag{7}$$

In this formulation, the objective function Eq. 2 is to assign crowd-shippers to satisfy as many online orders as possible. The second part of the objective function in Eq. 3 minimize the sum of extra traveling distance rate for all crowd-shippers; The two parts are weighted via the coefficients ω_1 and ω_2, for reducing the influence made by the second part; The equation 4 defines the extra travel distance for a crowd-shipper i delivery online order j; The constraint 5 guarantees that each crowd-shipper can only delivery no more than one online order; constraint 6 ensures one online order can only be delivered by one crowd-shipper; Finally, the constraints in Eq. 7 define the problem variables.

Solving the optimization problem determines the optimal crowdshipper-order assignments, reported by a mapping $\mu : \hat{\mathcal{I}} \longrightarrow \hat{J}$, the set of $\hat{\mathcal{I}} \subseteq \mathcal{I}$ of the crowd-shippers to actually delivery orders. In addition, set of $\hat{J}$ of the orders to actually be delivered by crowd-shippers. For example, if an order $j \in \hat{J}$ is determined to be delivery by a crowd-shipper $i \in \hat{\mathcal{I}}$, then $\mu(i) = j$.

4.2. Second stage optimization

In the second stage optimization, an optimal compensation scheme is determined with considering the following points:

- crowd-shipping increases the probabilities that online orders are satisfied, with a consequent increase in profits.
- if crowd-shippers accept assigned online orders, the number of orders delivered by professional drivers decreases with a consequent staff cost reduction.
- the second consideration holds only if the given compensations are smaller than the saved, delivered costs.

Since the second optimization problem's objective is to determine the set of optimal compensation scheme that maximizes the saving of the store. In doing so, for each crowd-shipper, we need to analyze his or her alternatives, such that:

- a: the crowd-shipper accepts to delivery online order;
- n: the crowd-shipper rejects to delivery online order.

The utility function of the two alternatives of the crowd-shipper is defined as:

$$\begin{cases} \boldsymbol{U}^a = \boldsymbol{V}^a + \boldsymbol{\epsilon}^a \\ \boldsymbol{U}^n = \boldsymbol{V}^n + \boldsymbol{\epsilon}^n \end{cases} \tag{8}$$

being $\boldsymbol{U}^a$ and $\boldsymbol{U}^n$ the vector gathering the perceived utilities of the two alternatives for all the crowd-shippers $i \in \hat{I}$, and $\boldsymbol{\epsilon}^a$ and $\boldsymbol{\epsilon}^n$ the vectors gathering the relevant stochastic residuals. Furthermore, $\boldsymbol{V}^a$ and $\boldsymbol{V}^n$ are the vectors of deterministic systematic utilities, here, we assume that $\boldsymbol{V}_i^a$ is dependent to the proposed compensation s_i while $\boldsymbol{V}^n$ is equal to the linear combinations of the extra distance $\mathcal{D}_{i,\mu(i)}$ and the inertia of crowd-shipper i. We defined the deterministic parts of the utility functions, inspired from [13] as:

$$\begin{cases} V_i^a = \beta_s s_i \\ V_i^n = \beta_{\mathcal{D}} \mathcal{D}_{i,\mu(i)} + \beta_{\mathcal{B}_v} \end{cases} \tag{9}$$

where, $\beta_{\mathcal{B}_v} \geq 0$ is a factor that takes into account the inertia of the customer. And the parameters β_s and $\beta_{\mathcal{D}}$ are suitable coefficients to be estimated. In addition, we model the residuals ϵ_i^a and ϵ_i^n, $\forall i \in \hat{I}$ as independent and identically distributed

Gumbel stochastic variables with null expectation and variance $Var[\epsilon_i] = \pi^2/6$, the probability that crowd-shipper accepts assigned task is defined as:

$$p_i^a = \frac{1}{1 + exp(V_i^n - V_i^a)} \tag{10}$$

whereas the probability that the crowd-shipper i reject to delivery order is defined as $p_i^n = 1 - p_i^a$. The optimization model of the problem is defined as:

$$\max_{\boldsymbol{s}} \sum_{\forall i \in \hat{\mathcal{I}}} (c_{\mu(i)} - s_i) p_i^a \tag{11}$$

subject to:

$$c_{\mu(i)} = c_0 + \alpha_c \mathcal{D}(store, d_j) \tag{12}$$

$$s_i \leq \omega_3 c_{\mu(i)}, \ \ \forall i \in \hat{\mathcal{I}} \tag{13}$$

$$s_i \geq 0, \ \ \forall i \in \hat{\mathcal{I}} \tag{14}$$

In the formulation, the objective function 11 maximizes the total cost reduction of the store. Constraint 12 defines the original shipping cost of the order assigned to crowd-shipper i. Constraint 13 guarantees the value of compensation can not be over a percentage of the original cost of the order. While the Constraint 14 defines the decision variables. In this section, we present a simple example to show how the crowd-shipping system works and validate the feasibility of the proposed approach. Then we randomly generate six groups of instances and compute their delivery costs by using the proposed approach and compare the results with a conventional delivery system that uses professional drivers to deliver orders.

5. COMPUTATIONAL STUDY

In this section, we evaluate the cost-effectiveness of the proposed two-stage optimization framework through a computational study. Since there are no benchmark instances for the crowd-shipping system that uses in-store customers to deliver online orders, we conducted computational studies on randomly generated instances. We consider a crowd-shipping system that has a store and five serviced zones $Z = \{A, B, C, D, E\}$, the distances between zones and distances between zones and the store are shown in Table 3. And the model parameters are reported in Table. 4.

We randomly generate six groups of crowd-shippers and online orders. We assume that the arrival rates of in-store customers and online orders follow Poisson Distributions as $\lambda_1 = 60$ and $\lambda_2 = 100$ respectively. To evaluate the cost-effectiveness of the proposed crowd-shipping system, we compare three different same-day delivery strategies under the same market setting. The first strategy ($A1$) is the traditional delivery strategy that uses company-owned vehicles to deliver orders; the second strategy ($A2$) pays a fixed fee plus a proportional to extra travel distance to crowd-shippers; the third one ($A3$) is what we proposed in this paper, an optimal compensation scheme. Moreover, to distinguish from professional drivers, in the second strategy, we assume that the compensation paid to crowd-shipper is 3 dollars plus 0.15 dollars per mile. The obtained results are shown in Table. 5.

	Store	A	B	C	D	E
Store	0	3	9	7.6	14.8	6.6
A	3	0	7.6	9.6	15.1	7.9
B	9	7.6	0	16.2	24.6	12.8
C	7.6	9.6	16.2	0	10.4	10.5
D	14.8	15.1	24.6	10.4	0	16
E	6.6	7.9	12.8	10.5	16	0

Table 3. Distance between each zones and store (km)

Parameter	Value	Parameter	Value
ω_1	0.8	β_s	1.5
ω_2	0.2	$\beta_{\mathcal{D}}$	0.3
ω_3	0.9	β_{B_v}	0.2
α_c	0.3	c_0	3
μ	0.25 h	σ^2	0.0025

Table 4. Parameters Settings

The results shown that:

- Around half of the crowd-shippers can be selected as candidate shippers to deliver online orders after the calculation of the first stage optimization model;
- On average, the initial expected cost of delivery is reduced by 3.08% due to the adoption of strategy $A2$;
- On average, the initial expected cost of delivery is reduced by 7.30% due to the adoption of strategy $A3$;

Computational experiments were executed on a computer with an Intel Core i7 6-core CPU with 16 GB of RAM, running at 2.6 GHz, using Mac OS X version 11.0.1. The model was implemented in Python version 3.8.5, using Fujitsu Digital Annealer.

6. CONCLUSION AND FUTURE WORK

This paper's main contribution is proposing a dynamic crowd-shipping system that can repeatedly match online orders and crowd-shippers by considering the probability that crowd-shippers reject assigned delivery tasks. To solve the matching problem, we propose a two-stage stochastic optimization model, and we also introduce a DES simulation framework to study the influence of each crowd-shipper's behaviors on the whole system. Also, the proposed system has been proved as feasibility and cost-effective after conducting the computational experiments.

Group	$\|J\|$	$\|I\|$	$\|\hat{I}\|$	A_1	A_2	Saving of A_2	A_3	Saving of A_3
1	87	52	32	487.2	471.57	3.20 %	443.76	8.91 %
2	107	56	35	588.48	565.64	3.88%	541.18	8.04%
3	97	55	31	524.82	506.27	3.53%	488.53	6.91%
4	95	49	26	536.46	518.80	3.29%	495.84	7.57%
5	108	53	31	607.08	593.61	2.22%	570.90	5.96%
6	99	54	31	531.18	518.79	2.33%	497.15	6.41%
Average	100	60	31	545.87	529.11	3.08%	506.23	7.30%

Table 5. Results obtained

In this paper, we consider a one-to-one matching problem; nevertheless, some in-store customers may be able to deliver more than one online order. In this case, a one-to-more matching problem, followed by the associated routing problem, will be a possible extension of our work. In addition, we assume that the utility functions of in-store customers and distribution of market demand are known in advance. We can address these two issues more effectively by leveraging data analytics and machine learning methods in future works.

7. REFERENCES

[1] Katarzyna Gdowska, Ana Viana, and João Pedro Pedroso, "Stochastic last-mile delivery with crowdshipping," *Transportation research procedia*, vol. 30, pp. 90–100, 2018.

[2] A Barr and J Wohl, "Exclusive: Walmart may get customers to deliver packages to online buyers," *REUTERS–Business Week*, , no. March, 2013.

[3] Jean Francois Rougès and Benoit Montreuil, "Crowdsourcing delivery: New interconnected business models to reinvent delivery," in *1st international physical internet conference*, 2014, vol. 1, pp. 1–19.

[4] Greg Bensinger, "Amazon's next delivery drone: You," *Wall street journal*, vol. 108, 2015.

[5] Claudia Archetti, Martin Savelsbergh, and M Grazia Speranza, "The vehicle routing problem with occasional drivers," *European Journal of Operational Research*, vol. 254, no. 2, pp. 472–480, 2016.

[6] Nabin Kafle, Bo Zou, and Jane Lin, "Design and modeling of a crowdsource-enabled system for urban parcel relay and delivery," *Transportation research part B: methodological*, vol. 99, pp. 62–82, 2017.

[7] Yuan Wang, Dongxiang Zhang, Qing Liu, Fumin Shen, and Loo Hay Lee, "Towards enhancing the last-mile delivery: An effective crowd-tasking model with scalable solutions," *Transportation Research Part E: Logistics and Transportation Review*, vol. 93, pp. 279–293, 2016.

[8] Fangxin Wang, Yifei Zhu, Feng Wang, and Jiangchuan Liu, "Ridesharing as a service: Exploring crowdsourced connected vehicle information for intelligent package delivery," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. IEEE, 2018, pp. 1–10.

[9] Iman Dayarian and Martin Savelsbergh, "Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders," *Production and Operations Management*, vol. 29, no. 9, pp. 2153–2174, 2020.

[10] Alp M Arslan, Niels Agatz, Leo Kroon, and Rob Zuidwijk, "Crowdsourced delivery—a dynamic pickup and delivery problem with ad hoc drivers," *Transportation Science*, vol. 53, no. 1, pp. 222–235, 2019.

[11] David Soto Setzke, Christoph Pflügler, Maximilian Schreieck, Sven Fröhlich, Manuel Wiesche, and Helmut Krcmar, "Matching drivers and transportation requests in crowdsourced delivery systems," 2017.

[12] Chao Chen and Shenle Pan, "Using the crowd of taxis to last mile delivery in e-commerce: a methodological research," in *Service orientation in holonic and multi-agent manufacturing*, pp. 61–70. Springer, 2016.

[13] Angela Di Febbraro, Nicola Sacco, and Mahnam Saeednia, "One-way car-sharing profit maximization by means of user-based vehicle relocation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 628–641, 2018.

OBSERVATIONAL LEARNING: IMITATION THROUGH AN ADAPTIVE PROBABILISTIC APPROACH

Sheida Nozari[1,2], Lucio Marcenaro[1], David Martin[2] and Carlo Regazzoni[1]

Department of Engineering and Naval architecture (DITEN), University of Genoa, Italy[1]
Intelligent systems lab, University Carlos III de Madrid, Spain[2]

ABSTRACT

This paper proposes an adaptive method to enable imitation learning from expert demonstrations in a multi-agent context. The proposed system employs the inverse reinforcement learning method to a coupled Dynamic Bayesian Network to facilitate dynamic learning in an interactive system. This method studies the interaction at both discrete and continuous levels by identifying inter-relationships between the objects to facilitate the prediction of an expert agent. We evaluate the learning procedure in the scene of learner agent based on probabilistic reward function. Our goal is to estimate policies that predict matched trajectories with the observed one by minimizing the Kullback-Leiber divergence. The reward policies provide a probabilistic dynamic structure to minimise the abnormalities.

Index Terms— imitation learning, multi-agent learning, reinforcement learning, Dynamic Bayesian network, performance analysis

1. INTRODUCTION

Imitation learning (IL) [1] approaches aim to mimic an expert behavior by transferring skills through observations and by following the demonstrations step-by-step [2]. However, imitating each step often becomes impracticable when the learning-agent and the environment are different from those in the demonstration. Also, using IL to track and reach a target in motion is still a challenging task. In many cases, the agent does not have to follow the expert unconditionally. Instead, it must care about the demonstrator's intention or the goal-based imitation [3]. A moving object can be modeled as a series of interactions with its surroundings such that its dynamics result from forces that act on it over time [4].

Modeling and understanding expert demonstrations (e.g., trajectories) are essential tasks in the success of multi-agent learning in a dynamic environment such as intelligent transportation [5], autonomous systems [6–8] and sports tracking data [9].

In order for autonomous multi-agent to learn such skills, they need a supervision signal that indicates the goal of the expected behavior. Typically, this supervision can come from a reward function in reinforcement learning (RL) that specifies which states and actions are desirable [10]. Recent advances in RL have improved IL to learn complicated behaviors in dynamic environments [11]. The integration of both modalities, RL and IL, enables the learning of complex skills from raw sensory observations [12]. However, the reward function in RL is task-specific, and the difficulty of manually specifying a reward function represents a significant barrier to the broader applicability of RL in complex observations [13]. Inverse reinforcement learning (IRL) [14] bypasses this issue by assuming that an agent receives the sequences of observation-action tuples. It tries to learn how to map observations to actions from these sequences through estimating a reward function. By approximating this function rather than directly learning the state-action, the apprentice is able to learn a reward function in new scenarios that explains the observed expert behavior. Moreover, it allows adapting to perturbation in it allows adapting to perturbation in a dynamic environment [15]. Accordingly, the demonstrations can be explained by a set of configurations between the moving objects at each time instant. Therefore, the complex models are enable to explain the interaction between objects and their surroundings [16]. We aim to take advantage of such interactions in a probabilistic manner through a coupled Dynamic Bayesian Network (DBN) structure to dynamically estimate present and future states at continues and discrete level. DBNs have been used for representing temporal relationships of the agent and a dynamic target. It is the case of predictive models based on objects' locations and their time derivatives [17–20].

To build this interaction model, we first use a set of spatial zones in a scene where the configurations are valid based on multiple expert demonstrations. Then, the transitions between the zones are used to track observations by employing a set of Kalman filters (KF) [21] coupled with a Particle filter (PF) [22] method to take advantage of both discrete and continuous variables under an interaction assumption. Finally, we employ the IRL approach using Q-network [23] to extract the probabilistic reward function regarding the detected abnormalities to match and evaluate the learner agent state trajectory (evidence) with the expert's demonstration (expectation) (Fig.1). We employ simulated data to validate the proposed method performance at the interacting rules into probabilistic models.

Our contributions are summarized as follows: *i)* we employ IRL approach to a coupled DBN structure that facilitates the characterization of objects' dynamics and their inter-relationships; *ii)* we learn a probabilistic multiple reward functions without exploiting the expert demonstration explicitly; *iii)* inferences from the proposed integrated method are used to minimise the abnormalities depending on the state of their surroundings. Learning a probabilistic reward function allows us to take uncertainty about the agent's dynamic into account, which reduces learning bias due to model errors.

2. TRAJECTORY REPRESENTATION

Probabilistic Graphical Models (PGMs) employ graph-based representation to encoding a variety of multi-dimensional random variables and represent causal relationships among them [24]. A particular type of PGM is the Dynamic Bayesian Network (DBN) [25]. Due to its hierarchical nature, DBN can express the temporal relationship between high-level variables (capturing abstract semantic information of the world) and low-level distributions (capturing rough sensory information of the environment) with their respective evolution through time. Recent works studied several algorithms for

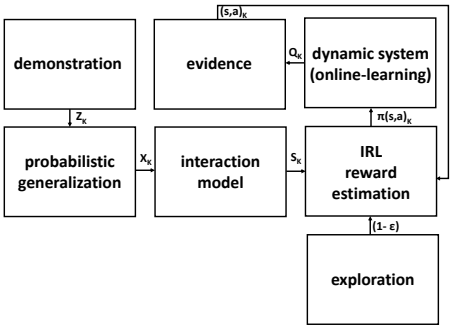


Fig. 1: Overview of the system.

inference in PGMs based on a data-driven way [26, 27]. A modern inference mechanism, namely, the Markov Jump Particle Filter (MJPF) presented in [26] can be employed to facilitate the generation of behavior based on DBN models learned computationally from data.

2.1. Dynamic Interaction Model

Let Z_k^1 and Z_k^2 be the observed positions of two entities, namely *Teacher (Te)* and *Reference Target (T_{Ref})*. Both agents are assumed to interact with each other at a given time instant k. Let us consider a KF which uses zero order motion dynamical equation:

$$\tilde{X}_k = A\tilde{X}_{k-1} + w_k, \tag{1}$$

where $\tilde{X}_k$ represents the object's state composed of its generalized coordinate positions and their velocities in a time instant k, such that $\tilde{X}_k = [\boldsymbol{x} \;\; \dot{\boldsymbol{x}}]^\intercal$ where $\boldsymbol{x} \in \mathbb{R}^d$ and $\dot{\boldsymbol{x}} \in \mathbb{R}^d$. d represents the number of coordinates of the environment. In (1), $A = [A_1 \;\; A_2]$ is a dynamic model matrix where $A_1 = [I_d \;\; 0_{d,d}]^\intercal$ and $A_2 = 0_{2d,d}$. I_n represents a square identity matrix of size n and $0_{l,m}$ is a $l \times m$ null matrix. w_k represents the prediction noise which is here assumed to be zero-mean Gaussian for all variables in X_k with a covariance matrix Q, such that $w_k \sim \mathcal{N}(0, Q)$. The proposed model in (1) suggests that moving objects will rest in a quasi-static location and only random noise perturbations, modeled by w_k will affect their states. At each time instant k a new measurement Z_k is made and it is assumed a linear relationship between Z_k and $\tilde{X}_k$, such that:

$$Z_k = H\tilde{X}_k + v_k, \tag{2}$$

where $H = [I_d \;\; 0_{d,d}]$ is the observation matrix that maps hidden states ($\tilde{X}_k$) to measurement (Z_k) and v_k is the measurement noise which is assumed to be zero-mean Gaussian with covariance R, such that, $v_k \sim \mathcal{N}(0, R)$. The deviations from predicted velocities are approximated using: $\dot{\boldsymbol{x}} = H^{-1}\big(Z_t - H\tilde{X}_{k-1}\big)$. A joint state space vector (System generalized states) is defined as $\tilde{\boldsymbol{X}}_k$ and consists of both *Te* and T_{Ref} states at each time instant k, such that:

$$\tilde{\boldsymbol{X}}_k = [\tilde{X}_k^1 \;\; \tilde{X}_k^2]^\intercal, \tag{3}$$

where $\tilde{X}_k^1$ and $\tilde{X}_k^2$ represent the Generalized States (GS) of Te and T_{Ref} respectively. To learn a situation model for our system, we perform a coupled DBN [16] by using two vocabularies, Te and T_{Ref}. The vocabularies are based on generalized joint states coming from training examples that describe a specific type of interaction between the objects. Each vocabulary is composed of configurations where $\tilde{\boldsymbol{X}}_k$ data is clustered. Each configuration represents a region where quasi-linear models are valid to present the interactive dynamical system over time (Fig.2.a). Vocabularies are defined as:

$$\boldsymbol{S}^i = \{s_1^i, s_2^i, \ldots, s_{L_i}^i\}, \tag{4}$$

where L_i is the total number of prototypes associated with the object i and s_l^i indexes the cluster of generalized joint states that favors object i's motion.

In a time instant k, each object i is represented by a situation state $S_k^i \in \boldsymbol{S}^i$. Active situation state from different objects are considered together as an activated configuration. For our case, the activated configuration at time instant k is written as $D_k = [S_k^T e, S_k^{T_{Ref}}]^\intercal$. Consequently, it is possible to define a dictionary containing possible configuration, such that:

$$\boldsymbol{D} = \{D^1, D^2, \cdots, D^M\}, \tag{5}$$

where D^m encodes a given identified configuration, M represents the total number of configurations (situation states combinations) and $D_k \in \boldsymbol{D}$. So, the configurations are created based on the different situation states related to the considered objects at the same time instant. Thus, $\boldsymbol{D}$ defines the whole system's discretization and the corresponding dynamics.

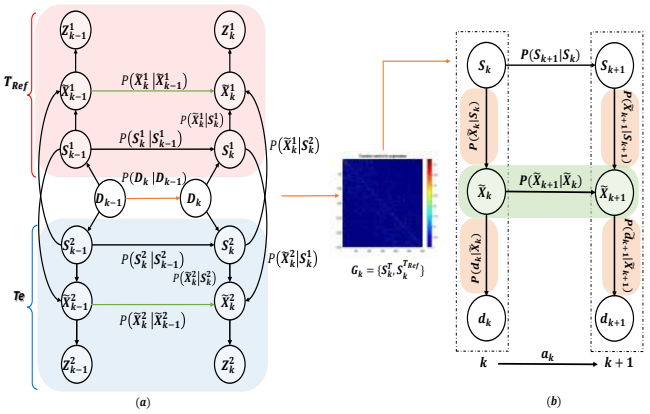


Fig. 2: a) A coupled-DBN for the interaction between Te and T_{Ref}. b) The provided learning model by L.

Transition model at the discrete level. By observing the configurations over time, it is possible to estimate a set of temporal transition matrices that encode the probabilities of passing from a current configuration to another one to estimate $p(D_k|D_{k-1}, t_k)$, where t_k encodes the time spent in the current word D_{k-1}.

Linear dynamic model at the continuous level. The object's motion can be modeled based on quasi-constant velocity, that is a function of the previously obtained regions S^i.

2.2. Probabilistic Learning Model

Each situation configuration S^i includes *Te* and T_{Ref} features $[(X_{bi}, V_{bi}), (X_{\beta j}, V_{\beta j})]_s$, where X_{bi} and $X_{\beta j}$ represent the position of *Te* and T_{Ref}, and V_{bi} and $V_{\beta j}$ represent the velocity of *Te* and T_{Ref}. Here, we can associate an average distance d_z to each configuration, that is a difference between X_{bi} and $X_{\beta j}$.
The position data is not meaningful because the agent in a dynamic environment usually deal with limited information. In order, by moving from one reference configuration to the other one, the system computes the distance at each time instant. This feature must

be comparable with the current model. Current model is based on the Learner agent *(L)* and the current target *(T_{Cur})* in the real-time through the online learning. Also, in the current model, we consider the interaction between *L* and T_{Cur} as a configuration at each time instant $[(X_{bi}, V_{bi}), (X_{\beta j}, V_{\beta j})]_{cur}$. Therefore, *L* measures the distance from the target which will change each time due to the action performed.

L uses the transition model estimated from the situation model (i.e. by observing the interaction between *Te* and T_{Ref}) to learn a new DBN encoding the dynamic behaviour followed by the teacher to reach the dynamic target (Fig.2-b). The transition model encodes the probability of moving from a certain configuration $[(X_{bi}, V_{bi}), (X_{\beta j}, V_{\beta j})]_{k-1}$ to another one $[(X_{bi}, V_{bi}), (X_{\beta j}, V_{\beta j})]_k$ in the situation model. In this way, *L* can predict the expected future configurations based on the dynamic transition rules encoded in the model and imitates a similar trajectory as the one it observed from *Te*.

We employ the MJPF [28] which uses a combination of PF and KF for prediction and inference purposes. Using the MJPF allows to predict the interaction among configurations at different levels: i) at the discrete level, to predict future configurations by means of PF which uses the transition probabilities ($p(D_k|D_{k-1})$) encoded in the transition model as a proposal distribution to propagate a set of particles realizing the predicted discrete variables (i.e. configurations); ii) at the continuous level, where velocity measurements and motion estimation of the states are are predicted using a bank of KFs.

Both levels provide a qualitative comparison between the current model's evidence in real-time and the corresponding prediction of the situation model through the learning procedure. Belief in hidden variables can be updated after receiving a new observation. Here the observations are the estimated distances $d(Te, T_{Ref})$. Then we estimate the expected d_z in the next time instant.

3. LEARNING DYNAMIC MULTIPLE REWARD

The objective of learning the reward policies is to integrate the IL with IRL by taking turns to i) optimize imitation policies that minimise the abnormalities (imitation loss). Hence, here, learning is relatively robust to modeling errors. ii) provide a probabilistic dynamic structure by an interactive reward estimation.

This work hypothesizes that during the learning phase, the learner uses a probabilistic interactively model. It employs the model in a Q-network [23] context for *i)* learning a multiple reward function and *ii)* regulating the learners' movement in the learning phase. We explain both contributions as below.

3.1. Reward function

Two different policies are considered:

Policy I. Learning to minimise the difference between the current learner's action, a_k and the mean action $\dot{\mathcal{C}}_{k-1}$ of the activated configuration $S_k \in \boldsymbol{\mathcal{S}}_{train}$ in the situation model, such that:

$$P_k^I = d_{\mathcal{M}}\big(g_a(S_k),\ a_k\big), \tag{6}$$

where $g_a(\cdot)$ is a function that extracts the action-distribution from a GS-distribution, such that $g_a(S_k) \sim \mathcal{N}(\dot{\mathcal{C}}_k,\ \Sigma_k^a)$ and Σ_k^a is the action's covariance information. $d_{\mathcal{M}}(\mathtt{X}, \mathtt{x})$ is the Mahalanobis distance [29] between a distribution $\mathtt{X}$ and a point $\mathtt{x}$. $S_k \sim \mathcal{N}(\tilde{\mathcal{C}}_k,\ \Sigma_K)$, which can be written as:

$$S_k = \underset{\mathcal{S}_m}{\operatorname{argmin}} \|s_k - \mathcal{C}_m\|_2. \tag{7}$$

Policy II. Learning to minimise the divergence between the distribution over the learner state (S_k) (calculated after taking an action a_{k-1}) and the discrete probability $p(S_k|S_{k-1})$ from the situation model (calculated by transition model $(d_z)_k$). The term S_{k-1}, required in $p(S_k|S_{k-1})$, is calculated based on Eq.(7).

The PF is employed to provide distributions over the learner state to have dynamic weight in the reward computation. The goal is to track the distributed state sequence (P_k) of a dynamic model. The distributed state emphasizes imperfect measurement from the current model by adapting noise to the learner state. The probability distribution over the learner state allows us to represent the uncertainty about the agent's dynamics.

For estimating d, two sources of information are required, the prior knowledge on how the d_k is expected to evolve and a measurement model related to evaluated (P_k). Here, we use the transition model to find the expected d, and we calculate Kullback–Leibler (KL) divergence [30] between two estimations, the d_z and d_{pi} to adjust the learner state. The KL presents a control input on the particles' weight. KL is used to refine the particles by comparing the expectation and the current model's measurements. The particles with the higher likelihood survive, and we use the mean of them to have probabilistic reward by considering the uncertainty. The policy II can be written as:

$$P_k^{II} = d_{\mathcal{M}}\big(g_s(d_{k|k-1}) \mid\mid \bar{X} \sum_{i=1}^{n} d_{pi}\big), \tag{8}$$

where $g_s(\cdot)$ as a function that extracts the state-distribution from a GS-distribution, such that $g_s(S_k) \sim \mathcal{N}(\mathcal{C}_k,\ \Sigma_k^s)$. Σ_k^s is the state's covariance information.

This paper considers both policies in parallel as a reward:

$$R_k := P_k^I + P_k^{II}. \tag{9}$$

3.2. Abnormality measurement

This work proposes an abnormality measurement based on the KL [30] divergence between the situation states $p(\tilde{S}_k^i|\tilde{S}_{k-1}^i)$ and the evidence $p(d_k|\tilde{X}_k)$, to evaluate the povided current model, such that:

$$\lambda_k^i = \int p(\tilde{S}_k^i|\tilde{S}_{k-1}^i) log \frac{p(\tilde{S}_k^i|\tilde{S}_{k-1}^i)}{p(d_k|\tilde{X}_k)}. \tag{10}$$

The values of λ_k^i indicates how much the prediction is supported by the observation. If the observation matches the prediction, then λ_k^i is close to 0. Otherwise the prediction deviates from the observation which leads to a high value of λ_k^i (close to 1) revealing the presence of an abnormality.

4. RESULTS

In this section, we provide numerical results to validate the proposed method. We consider a table of trained data where L, chases T_{Cur} in a 40×40 space Fig.3. In training data, L's motion is described by 8 different motion unit-vectors associated with the cardinal and intercardinal directions. T_{Cur} motions consists in a horizontal dynamics along the x axis at a fixed height point $y_{T_{Cur}}$. Accordingly, T_{Cur} can move in two senses: right or left inside the interval $[x_{T_{Cur}}^{(min)}, x_{T_{Cur}}^{(max)}]$. T_{Cur} dynamics consists of a continuous motion in one sense until it reaches an interval boundary. Then, it starts moving in the opposite sense covering only the defined interval points. The speed of T_{Cur} movement is different from T_{Ref} in the situation

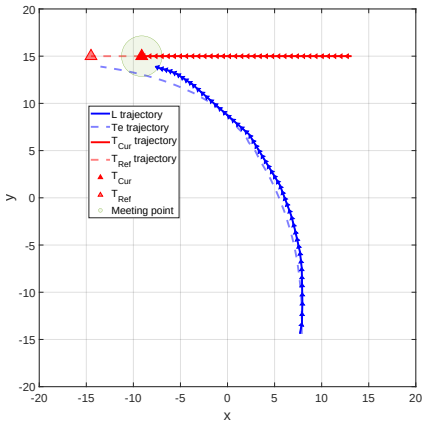


Fig. 3: Example of matched trajectories.

model to guarantees that L learns to reach the target in a new scenario. The following parameters are employed for simulation purposes: $y_{T_{Cur}} = 15$, $x^{(min)}_{T_{Cur}} = -15$ and $x^{(max)}_{T_{Cur}} = 15$. Results related to the capabilities of detecting abnormalities and evaluating the current model are explained in detail as follows.

Abnormality detection. Evaluating the current model's configurations during the learning phase is employed to detect abnormalities. Training includes 500 episodes from different start positions. In each episode L is trained to reach T_{Cur} 8 times. It means L tries 4000 trajectories through 500 different start positions. Fig. 4 shows the result of motion's difference between L and Te at the continuous level at time k by using Mahalanobis distance [29].

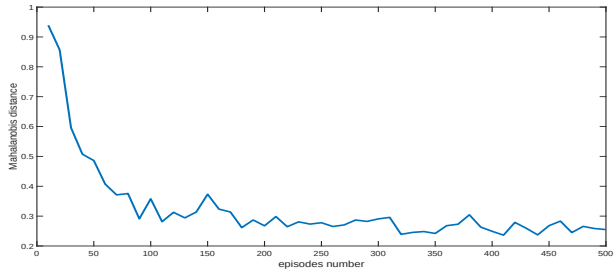


Fig. 4: Motion difference.

Fig. 5 shows abnormality estimation based on the divergence between the current model's configuration and the situation model's prediction at the discrete level at time $k+1$ through KL divergence measurements.

From both figures, it is possible to see how high abnormality values are present in the learning's initial portion. Once L learns the reward policies, the measurements go down dramatically. In Fig. 5, although the divergence measurements is not too high (the highest value is 11×10^{-3}), L learns to minimise it.

Current model evaluation. Here the situation model is available as a ground truth. To evaluate the current model's efficiency, we translate the testing phase's result to a switching DBN based on L and T_{Cur} interaction. Fig. 6 shows the result of a comparison between the motions generated based on the situation model and the respective evidence of the translated current model by using KL measurements.

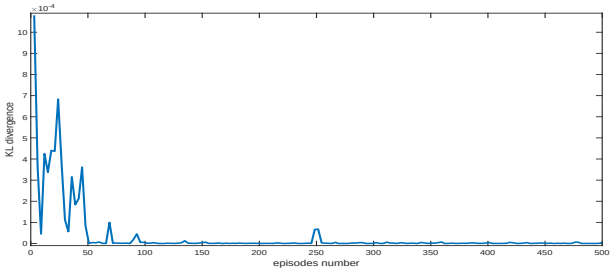


Fig. 5: Convergence measurement.

As Fig. 6 shows, when L's distance to T_{Cur} is between [15,40], where L follows the expert trajectory, the abnormality estimation is lower than other positions. In the range [10,15], the measurement increases gradually because L tries to adapt to T_{Cur} behavior, which is different with T_{Ref}. The highest difference belongs to the distance between [0, 5] to the target, where L's motion is goal-based. However, most of the abnormality measurements (75%) are less than 0.03, that as we mentioned previously, values close to 0 indicate that evidence matches with the expectation. As the results show, the learner agent succeeded in imitating the teacher's behavior.

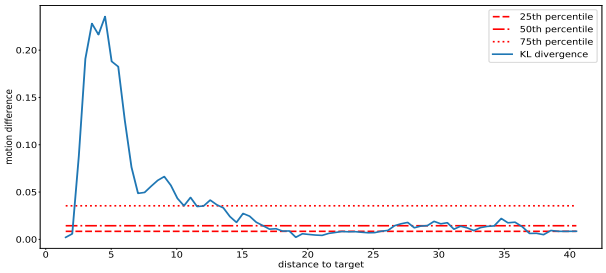


Fig. 6: Motion difference based on distance from target.

5. CONCLUSION

In this paper, we proposed an adaptive probabilistic model for IL. Algorithms for performing inferences and learning the probabilistic reward structure are presented, which enables the learning-agent to take uncertainty appropriately into account. Our method demonstrates learning from an interaction model to estimate the reward function through online learning. Experimental results show the capability to minimise the abnormalities while learning the policies from the demonstrations. Those abnormalities can be used as qualitative observation from expert demonstrations in order to learn from unseen configurations. Comparisons between the simulated learner agent and encoded DBN configurations in the proposed model can encode multiple IRL policies. Future works include more complex interactions between objects, such as multiple learner agents, to create robust DBN structures for IRL.

6. REFERENCES

[1] Stefan Schaal et al., "Learning from demonstration," *Advances in neural information processing systems*, pp. 1040–1046, 1997.

[2] Stefan Schaal, Auke Ijspeert, and Aude Billard, "Computational approaches to motor learning by imitation," *Philosophi-*

cal Transactions of the Royal Society of London. Series B: Biological Sciences, vol. 358, no. 1431, pp. 537–547, 2003.

[3] Deepak Verma and Rajesh PN Rao, "Goal-based imitation as probabilistic inference over graphical models," in *Advances in neural information processing systems*. Citeseer, 2006, pp. 1393–1400.

[4] Damian Campo, Alejandro Betancourt, Lucio Marcenaro, and Carlo Regazzoni, "Static force field representation of environments based on agents' nonlinear motions," *EURASIP Journal on Advances in Signal Processing*, vol. 2017, no. 1, pp. 1–15, 2017.

[5] Weiwei Jiang, Jing Lian, Max Shen, and Lin Zhang, "A multi-period analysis of taxi drivers' behaviors based on gps trajectories," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017, pp. 1–6.

[6] Mohamad Baydoun, Mahdyar Ravanbakhsh, Damian Campo, Pablo Marin, David Martin, Lucio Marcenaro, Andrea Cavallaro, and Carlo S Regazzoni, "A multi-perspective approach to anomaly detection for self-aware embodied agents," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 6598–6602.

[7] Wonteak Lim, Seongjin Lee, Myoungho Sunwoo, and Kichun Jo, "Hierarchical trajectory planning of an autonomous car based on the integration of a sampling and an optimization method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 2, pp. 613–626, 2018.

[8] Dennis Fassbender, Benjamin C Heinrich, Thorsten Luettel, and Hans-Joachim Wuensche, "An optimization approach to trajectory generation for autonomous vehicle following," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 3675–3680.

[9] Alina Bialkowski, Patrick Lucey, Peter Carr, Yisong Yue, and Iain Matthews, "Win at home and draw away: Automatic formation analysis highlighting the differences in home and away team behaviors," in *Proceedings of 8th annual MIT sloan sports analytics conference*. Citeseer, 2014, pp. 1–7.

[10] Richard S Sutton and Andrew G Barto, *Reinforcement learning: An introduction*, MIT press, 2018.

[11] Yuke Zhu, Ziyu Wang, Josh Merel, Andrei Rusu, Tom Erez, Serkan Cabi, Saran Tunyasuvunakool, János Kramár, Raia Hadsell, Nando de Freitas, et al., "Reinforcement and imitation learning for diverse visuomotor skills," *arXiv preprint arXiv:1802.09564*, 2018.

[12] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[13] Ashley Edwards, Charles Isbell, and Atsuo Takanishi, "Perceptual reward functions," *arXiv preprint arXiv:1608.03824*, 2016.

[14] Andrew Y Ng, Stuart J Russell, et al., "Algorithms for inverse reinforcement learning.," in *Icml*, 2000, vol. 1, p. 2.

[15] Thibaut Munzer, Bilal Piot, Matthieu Geist, Olivier Pietquin, and Manuel Lopes, "Inverse reinforcement learning in relational domains," in *International Joint Conferences on Artificial Intelligence*, 2015.

[16] Mohamad Baydoun, Damian Campo, Divya Kanapram, Lucio Marcenaro, and Carlo S Regazzoni, "Prediction of multi-target dynamics using discrete descriptors: an interactive approach," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 3342–3346.

[17] Guotao Xie, Hongbo Gao, Lijun Qian, Bin Huang, Keqiang Li, and Jianqiang Wang, "Vehicle trajectory prediction by integrating physics-and maneuver-based approaches using interactive multiple models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5999–6008, 2017.

[18] Yu Feng, Jian Sun, and Peng Chen, "Vehicle trajectory reconstruction using automatic vehicle identification and traffic count data," *Journal of advanced transportation*, vol. 49, no. 2, pp. 174–194, 2015.

[19] Xing Sun, Nelson HC Yung, and Edmund Y Lam, "Unsupervised tracking with the doubly stochastic dirichlet process mixture model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2594–2599, 2016.

[20] Jacinto C Nascimento, Mário AT Figueiredo, and Jorge S Marques, "Activity recognition using a mixture of vector fields," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1712–1725, 2012.

[21] Greg Welch, Gary Bishop, et al., "An introduction to the kalman filter," 1995.

[22] Fredrik Gustafsson, Fredrik Gunnarsson, Niclas Bergman, Urban Forssell, Jonas Jansson, Rickard Karlsson, and P-J Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Transactions on signal processing*, vol. 50, no. 2, pp. 425–437, 2002.

[23] Christopher JCH Watkins and Peter Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[24] Luis Enrique Sucar, "Probabilistic graphical models," *Advances in Computer Vision and Pattern Recognition. London: Springer London. doi*, vol. 10, pp. 978–1, 2015.

[25] Zoubin Ghahramani, "Learning dynamic bayesian networks," in *International School on Neural Networks, Initiated by IIASS and EMFCSC*. Springer, 1997, pp. 168–197.

[26] M. Baydoun, D. Campo, V. Sanguineti, L. Marcenaro, A. Cavallaro, and C. Regazzoni, "Learning switching models for abnormality detection for autonomous driving," in *2018 21st International Conference on Information Fusion (FUSION)*, July 2018, pp. 2606–2613.

[27] Yajing Zheng, Shanshan Jia, Zhaofei Yu, Tiejun Huang, Jian K Liu, and Yonghong Tian, "Probabilistic inference of binary markov random fields in spiking neural networks through mean-field approximation," *Neural Networks*, 2020.

[28] M. Baydoun, D. Campo, V. Sanguineti, L. Marcenaro, A. Cavallaro, and C. Regazzoni, "Learning switching models for abnormality detection for autonomous driving," in *2018 21st International Conference on Information Fusion (FUSION)*, 2018, pp. 2606–2613.

[29] Roy De Maesschalck, Delphine Jouan-Rimbaud, and Désiré L Massart, "The mahalanobis distance," *Chemometrics and intelligent laboratory systems*, vol. 50, no. 1, pp. 1–18, 2000.

[30] Solomon Kullback and Richard A Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.

DETECTING ANOMALOUS SWARMING AGENTS WITH GRAPH SIGNAL PROCESSING

Kevin Schultz, Anshu Saksena, Elizabeth P. Reilly, Rahul Hingorani, Marisel Villafañe-Delgado

Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Road, Laurel, MD 20723

ABSTRACT

Collective motion among biological organisms such as insects, fish, and birds has motivated considerable interest not only in biology but also in distributed robotic systems. In a robotic or biological swarm, anomalous agents (whether malfunctioning or nefarious) behave differently than the normal agents and attempt to hide in the "chaos" of the swarm. By defining a graph structure between agents in a swarm, we can treat the agents' properties as a graph signal and use tools from the field of graph signal processing to understand local and global swarm properties. Here, we leverage this idea to show that anomalous agents can be effectively detected using their impacts on the graph Fourier structure of the swarm.

Index Terms— swarming, graph signal processing, anomaly detection

1. INTRODUCTION

Collective motion in biological systems such as insect swarms, fish schools, and bird flocks are visually striking emergent behaviors that have motivated considerable research in biology, physics, and engineering [1–5]. Due to the distributed nature of swarming systems, graph theory has found considerable utility in the analysis and synthesis of swarming systems by modeling the communications or other interactions between swarming agents as a graph structure [6, 7]. Recently, tools from computational topology have been applied to understanding the structure of swarms in terms of connected subcomponents as well as the presence of holes and voids [8, 9]. These tools explicitly rely on the parametric construction of a graph structure between agents.

This topological approach was further extended in [10] to analyze how the local and global "order" of swarm properties varies with respect to the graph structure. The analysis in [10] employed the field of graph signal processing (GSP) and graph Fourier analysis to show that common swarm states were highly structured (i.e., band-limited) when viewed in the graph Fourier domain. Broadly speaking, GSP builds on its roots in spectral graph theory [11] and algebraic signal processing [12] to generalize concepts from classical signal processing to signals defined on the vertices of irregular domains modeled by graphs [13–16].

Anomaly detection [17] is among the application areas considered in the seminal GSP works [15, 18]. Since the initial work that considered temperature sensor networks [15] and more generally abstract sensor networks [18], GSP-based anomaly detection has been applied to a number of areas, including power systems [19,20], social networks [16], and image processing [21]. At a high level, these techniques exploit (generally low-pass) structure in the graphical Fourier transform (GFT) of some signal defined on the vertices of a graph, and then threshold on the signal content after (graph) filtering to remove this structure [16]. We note that this class of problem is fundamentally different from detecting an anomalous graph structure in a network, itself a well studied problem [22].

In this work, we show how the GFT structure of swarms revealed in [10] can be used to design "graph filters" that operate on the swarm state to detect agents within the swarm whose dynamics (and thus behavior) are fundamentally different from the bulk of the swarm. These anomalous agents have behaviors that differ in subtle ways from the rest of the swarm; from a purely kinematic perspective the anomalous trajectories are consistent with the non-anomalous agents. Instead, these behaviors are detected through interaction and comparison with their neighbors that manifest in the GFT domain as outliers. To our knowledge, this work is the initial adaptation of GSP techniques to the swarming domain, and the detection problems herein are challenging enough that multiple swarm measurements are needed for effective detection, demonstrating an anomaly detection problem where the signal is not only time-varying as in [23], but with a time varying graph, as well. This work is related to, but distinct from, the inference of dynamical parameters [24] and the identification of collective states [8,9,25], as this work more closely resembles clustering (i.e., unsupervised learning) of distinct behavior regimes within the swarm. More generally, this work falls under fault detection in swarms [26], and addresses similar problems as [27], which uses neural networks to identify joint collective anomalies, comprising a more data intensive approach.

In the following, we first review some GSP and swarming preliminaries. Next, we briefly discuss how to interpret

This work was partially supported by NSF award NCS/FO 1835279 and JHU/APL IR&D.

swarming in a GSP context and define two GSP approaches to the detection of anomalous agents in a swarm. We then consider both approaches in a range of scenarios using swarming simulations of [1] and [4], discussing how to leverage their unique graph Fourier signatures, and analyze the resulting anomaly detectors using detection theory. We conclude with summary remarks and discuss future directions.

2. GSP BACKGROUND

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a collection of N vertices $\mathcal{V} = \{v_i\}_{i=1}^N$ and edges $e_{ij} \in \mathcal{E}$ connecting nodes v_i and v_j. The graph adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ mathematically represents interactions in a graph, with nonzero entries A_{ij} indicating the presence of an edge e_{ij}. In this work, we are concerned with non-negative weighted undirected graphs, so $A_{ij} = A_{ji} \geq 0$. The degree matrix $\mathbf{D}$ is a diagonal matrix whose entries account for the total number of connections for each node and is defined as $D_{ii} = \sum_j A_{ij}$. The (combinatorial) graph Laplacian is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. A related matrix, the normalized Laplacian, is defined for connected graphs by $\bar{\mathbf{L}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$ and extends to general graphs several nice properties of the Laplacian that only hold for certain regular graphs [11].

GSP is concerned with the analysis of signals or functions defined on the vertices of a graph. Let $\mathbf{f} : \mathcal{V} \to \mathbb{F}^m$ be a so-called graph signal defined on $\mathcal{V}$ that takes values in some finite dimensional Hilbert space. We adopt the shorthand convention that $\mathbf{f}_i = \mathbf{f}(v_i)$ for the ith vertex of $\mathcal{G}$. Over the last decade, most GSP efforts have focused on extending techniques defined in classical signal processing to signals defined over graphs, such as filtering and multiple signal transformations including the GFT. In a natural definition for non-negative weighted symmetric graphs, the GFT uses the eigenvectors of $\mathbf{L}$ as basis functions instead of the complex exponentials used in the Fourier transform. However, [10] demonstrated that $\bar{\mathbf{L}}$ may be better suited for analysis of swarms. Using the eigendecomposition $\bar{\mathbf{L}} = \mathbf{U}\Lambda\mathbf{U}^\top$, with $\mathbf{U}$ a unitary matrix of eigenvectors, and Λ a diagonal matrix of eigenvalues in increasing order, the GFT of a graph signal $\mathbf{f}$ is defined as $\hat{\mathbf{f}} = \mathbf{U}^\top\mathbf{f}$ and the corresponding inverse GFT as $\mathbf{f} = \mathbf{U}\hat{\mathbf{f}}$. Using this definition of the GFT, the eigenvalues λ_i are a natural generalization of frequency that are no longer evenly spaced in general, but will be in the interval $[0, 2]$. This further suggests an intuitive mechanism to define graph filters using the eigendecomposition. Let $\mathbf{H}$ be a diagonal matrix, then $\mathbf{U}\mathbf{H}\mathbf{U}^\top\mathbf{f}$ is a filtered graph signal where the filter has "frequency response" H_{ii} at frequency λ_i.

3. SWARM MODELS

In this work, we will consider anomalous agents in an otherwise homogeneous swarm using two different swarming models: the biologically inspired model of [1] and the swarmalator model of [4]. The model in [1] uses disjoint behavior regions for repulsion, alignment, and attraction. Depending on the relative radii of these regions, the angle of a blind-spot behind each agent, and the amount of noise, the swarm dynamics will produce one of four steady states. In the absence of an alignment region, the dynamics produce a disorganized "swarming" state characterized by low local and global alignment in velocity and minimal collective displacement. When the region of alignment is larger than the region of repulsion, but still small relative to the region of attraction, the swarm tends to form three-dimensional torus-like structures, with a high level of local alignment in velocity but overall low global alignment and little collective displacement. As the alignment region approaches the attraction region, two additional collective states form, both with higher global velocity alignment and collective displacement than the swarming and torus states. Here, we focus on the first two states due to their apparent disorganization, presumably making an anomalous agent harder to detect (see Fig. 1).

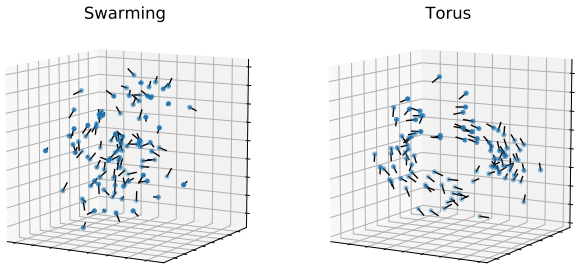


Fig. 1. Example swarm states for the model of [1]. Dots indicate agent positions and lines indicate velocities.

The swarmalator model was developed to understand joint swarming and synchronization behaviors, combining spatial attraction and repulsion with an auxiliary phase θ_j. This phase can modulate the sign of the spatial attraction/repulsion and is itself governed by dynamics in the vein of the Kuramoto model [28]. Of the several steady states observed in [4] we focus on the active wave state in two dimensions, where the agents form counter-rotating (roughly) phase-ordered rings that exhibit considerable dynamism in both the position and phase states (see Fig. 2, left panel).

4. METHODS

The necessary ingredients for performing GSP analysis are the specification of a graph and a function defined on the vertices of that graph. For a collection of N swarming entities with positions $\mathbf{x}_j$ and velocities $\mathbf{v}_j$ both $\in \mathbb{R}^n$ (time indices surpressed), following [10] we define each agent as a vertex in a graph, and for each instance in time we construct an ad-

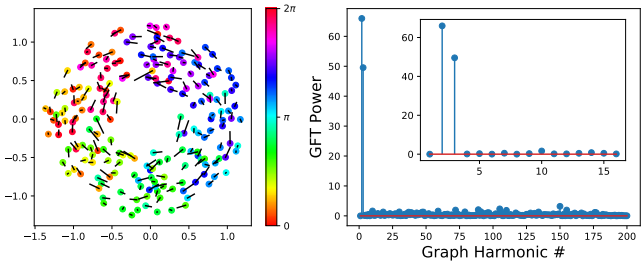


Fig. 2. Example swarmalator state. Left: Agent positions, with color indicating θ_j and lines indicating velocity. Right: GFT power using signal $\mathbf{h}$ for the swarm state on the left. Inset axes show zoomed view of spectral concentration.

jacency matrix $A_{jk} = \exp(-||\mathbf{x}_j - \mathbf{x_k}||_2^2/\sigma^2)$ ($j \neq k$, otherwise $A_{jj} = 0$), where $\sigma^2 = \frac{1}{N(N-1)} \sum_{j \neq k} ||\mathbf{x}_j - \mathbf{x}_k||_2^2$. Depending on the scenario, we consider different graph signals. For the model of [1] we consider the normalized velocity $\mathbf{u}_j = \mathbf{v}_j/||\mathbf{v}_j||_2^2$ and the adjusted position $\mathbf{r}_j = \mathbf{x}_j - \bar{\mathbf{x}}$ where $\bar{\mathbf{x}} = \frac{1}{N} \sum \mathbf{x}_j$. For the analysis of swarmalators we use the signal $\mathbf{h}_j = \exp(i\theta_j)$.

With these graph and graph signal definitions, using the GFT derived from $\bar{\mathbf{L}}$ we have from [10] that both the model of [1] and swarmalators exhibit spectral concentration in a few graph Fourier harmonics (see Fig. 3 and Fig. 2, right panel). Furthermore, these harmonics are low frequency, indicating local alignment of the graph signal with respect to the graph topology.

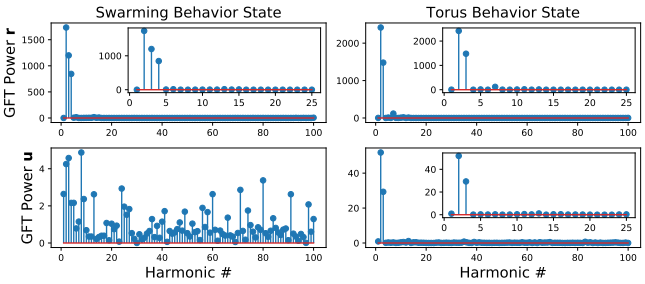


Fig. 3. Example GSP analysis of swarming (left) and torus (right) states from Fig. 1. (Top) GFT power of position states $\mathbf{r}$ and (Bottom) GFT power of normalized velocity states $\mathbf{u}$. Inset axes show zoomed view of spectral concentration.

This spectral concentration immediately suggests two approaches for anomaly detection: 1) a graph filter approach with filters defined by $H_{ii} = \mathbb{1}_S(i)$ where $\mathbb{1}_S(i) = 1$ if $i \in S$ and 0 otherwise and S is a set of graph frequencies expected not to be common in nominal swarms, and 2) using the intuition that an anomalous agent should be different in "smoothness" than the rest of the swarm, the high-pass filter $\mathbf{H} = \Lambda$, which has been suggested as a measure of local graph smoothness for a graph signal [16]. We call the first approach out-of-band power (OOBP) and the second local graph smoothness (LGS). In either case, the input graph signal $\mathbf{f}$ is filtered to generate $\mathbf{g} = \mathbf{UHU}^\top \mathbf{f}$, and $||\mathbf{g}_i||_2^2$ is used as a threshold statistic for the detection of an anomalous agent in the swarm (c.f., [18–20]).

5. RESULTS

We ran the swarming model of [1] with 100 agents (99 normal, 1 anomalous) and the swarmalator model with 200 agents (199 normal, and 1 anomalous). We ran each with a few different sets of parameters for both the nominal and anomalous agents. For the model of [1], we ran 4 cases:

- **Case 1**: Nominal behavior is Swarming, anomalous agent has slightly larger repulsion region ($\mathbf{r}$ is the graph signal)
- **Case 2**: Nominal behavior is Torus, anomalous agent has slightly larger repulsion region (same anomaly as Case 1, $\mathbf{r}$ is the graph signal)
- **Case 3**: Nominal behavior is Torus, anomalous agent has no alignment region ($\mathbf{u}$ is the graph signal)
- **Case 4**: Nominal behavior is Swarming, anomalous agent has large alignment region ($\mathbf{u}$ is the graph signal)

For the swarmalator model, the nominal behavior of positional attraction to like phases and phase repulsion to nearby phases is defined by parameters $A = B = J = 1$, $K = -0.75$ and the anomaly was parameterized at $J = -1$ for positional repulsion to like phases and various values of K in $[-0.75, 0.75]$ that varies the strength of the phase repulsion and attraction, using the model definitions as in [4, Fig. 4]. We use $\mathbf{h}$ as the graph signal, and we denote this **Case 5**.

We define OOBP filters for each case based on the nominal and anomalous behaviors. For both Case 1 and 2, the nominal $\mathbf{r}$ is lowpass (Fig. 3, top row) so we use the highpass filters $H_{ii} = \mathbb{1}_{\{5,\dots N\}}(i)$ and $H_{ii} = \mathbb{1}_{\{4,\dots,N\}}(i)$, respectively. For Case 3, $\mathbf{u}$ is lowpass (Fig. 3, bottom right) and we use the same filter as Case 2. Despite the bandlimited nature of these signals, we found that the exclusion of the first harmonic improved detection. For Case 4, we expect the nominal swarm state to be spectrally flat (Fig. 3, bottom left) and an anomaly with a large alignment range to contribute to low frequency components, so our OOBP filter should be lowpass. We find that $H_{ii} = \mathbb{1}_{\{1,\dots,6\}}(i)$ performed well. For Case 5, the signal $\mathbf{h}$ is bandlimited (Fig. 2, right), and the bandstop filter $H_{ii} = \mathbb{1}_{\{1\} \cup \{4,\dots,N\}}(i)$ produced excellent results.

While the LGS approach is already defined by a given swarm's graph, note that since anomalies in Case 4 are expected to appear in low frequencies, they will have low values in the LGS approach, but high values for the remaining cases. We set the direction of our detector accordingly.

Each of these model parameterizations was run with 100 Monte Carlo runs that varied the initial state of each agent and then iterated for 1500 time steps to allow for the swarm to

stabilize. After that, the graph was defined from agent states as described above and the two approaches OOBP and LGS were evaluated at discrete time steps 50 time steps apart to create swarm "snapshots" to allow for correlations in swarm state to die out. The results of the norm squared of the filtered signal at each agent were summed over the snapshots to produce the measurements that were thresholded to define our anomaly detector. The area under the receiver operating characteristic (ROC) curve, empirically estimating the probability that an anomalous agent whose parameters are uniformly randomly selected from those explored would have a higher statistic value than a randomly chosen nominal agent, providing an indication of how effectively the threshold statistic can be used to detect anomalous agents [29], is shown in Fig. 4.

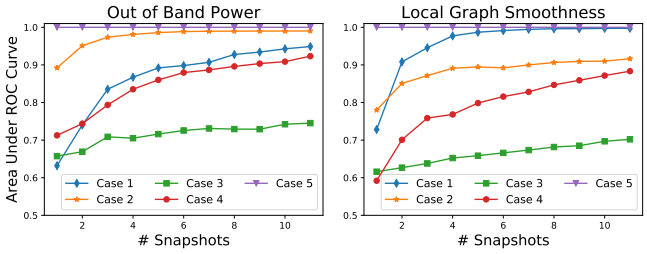


Fig. 4. Area under the ROC curves for the five anomaly cases.

These curves show that there is a range of efficacy of both detectors across the cases considered, although we note that all of them produce areas substantially greater than 0.5 (random chance), and all appear to improve as the number of snapshots (analogous to integration time) increases. It is clear that Case 5 appears to be an extremely easy detection problem, whereas Case 3 is the most challenging. Viewing a typical example from Case 5 (Fig. 5) we see that the anomalous agent is somewhat out of phase from the rest of the swarm, but may be challenging to visually identify if not explicitly marked. However, the filtered response is a clear outlier, consistent with the intuition behind our approach. The more challenging cases have considerable overlap in the distributions of the filtered signals. We note that detecting the repulsion-based anomaly in Cases 1 and 2 appears to be considerably easier than detecting anomalous alignment behaviors in Cases 3 and 4. Anecdotally we do see that more repulsive agents will occasionally appear farther from $\bar{\mathbf{x}}$ than nominal agents, whereas the alignment effects are generally unnoticeable visually. Another interesting facet of these results is that OOBP appears to be a more effective detection mechanism, except in Case 1. Why this is the case despite such a concentrated GFT response of the nominal behavior we leave to future research.

6. CONCLUSION

In summary, we demonstrated how the GFT structure of swarms could be leveraged to detect anomalous agent behaviors in a number of scenarios. We considered two different filter-based approaches and used a variety of graph signals as the subjects of the analysis. Overall, we demonstrated that the relative efficacy of these approaches appears highly dependent on the specific context. Using detection-theoretic ROC analysis we demonstrated that accumulating information over several snapshots of the swarming data may be needed to have effective detectors in the most challenging cases.

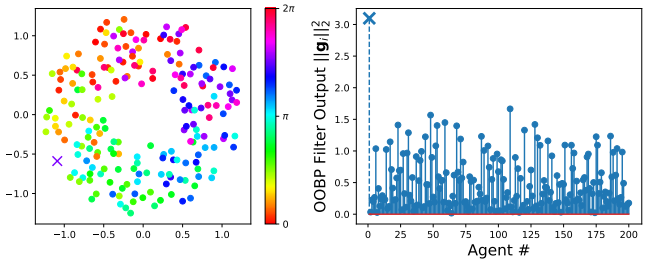


Fig. 5. Swarmalator state with anomalous agent, marked by $\times$ in both panes. Left: Agent positions, with color indicating θ_j. Right: Graph filtered response of $\mathbf{h}$.

This work indicates a number of potential future directions in not only anomaly detection in swarms but more generally in the application of GSP techniques. Further analysis of the cases considered here is warranted, and additional swarming models and anomalies could be analyzed as well. GSP techniques that analyze both the vertex and time domains simultaneously could be incorporated [21], although this would require handling time-varying graphs (rather than the independent manner in which they were treated here). In principle, this work could be combined with distributed graph filtering techniques to perform self-anomaly detection within the swarm. Finally, we note that the the techniques here could be adapted to behavior discrimination and classification of agents in a swarm such as leader vs. follower behaviors.

7. REFERENCES

[1] Iain D Couzin, Jens Krause, Richard James, Graeme D Ruxton, and Nigel R Franks, "Collective memory and spatial sorting in animal groups," *Journal of theoretical biology*, vol. 218, no. 1, pp. 1–12, 2002.

[2] David JT Sumpter, *Collective animal behavior*, Princeton University Press, 2010.

[3] Tamás Vicsek and Anna Zafeiris, "Collective motion," *Physics reports*, vol. 517, no. 3-4, pp. 71–140, 2012.

[4] Kevin P O'Keeffe, Hyunsuk Hong, and Steven H Strogatz, "Oscillators that sync and swarm," *Nature communications*, vol. 8, no. 1, pp. 1–13, 2017.

[5] Kevin M Passino, *Biomimicry for optimization, control,*

and automation, Springer Science & Business Media, 2005.

[6] Ali Jadbabaie, Jie Lin, and A Stephen Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on automatic control*, vol. 48, no. 6, pp. 988–1001, 2003.

[7] Reza Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on automatic control*, vol. 51, no. 3, pp. 401–420, 2006.

[8] Chad M Topaz, Lori Ziegelmeier, and Tom Halverson, "Topological data analysis of biological aggregation models," *PloS one*, vol. 10, no. 5, pp. e0126383, 2015.

[9] Padraig Corcoran and Christopher B Jones, "Modelling topological features of swarm behaviour in space and time with persistence landscapes," *IEEE Access*, vol. 5, pp. 18534–18544, 2017.

[10] Kevin Schultz, Marisel Villafañe Delgado, Elizabeth P Reilly, Anshu Saksena, and Grace M Hwang, "Analyzing collective motion using graph fourier analysis," *arXiv preprint: arXiv:2103.08583*, 2021.

[11] Fan Chung, *Spectral graph theory*, Number 92. American Mathematical Soc., 1997.

[12] Markus Puschel and José MF Moura, "Algebraic signal processing theory: Foundation and 1-d time," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3572–3585, 2008.

[13] David Shuman, Sunil Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 3, no. 30, pp. 83–98, 2013.

[14] Aliaksei Sandryhaila and José MF Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[15] Aliaksei Sandryhaila and Jose MF Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.

[16] Raksha Ramakrishna, Hoi To Wai, and Anna Scaglione, "A user guide to low-pass graph signal processing and its applications: Tools and applications," *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 74–85, 2020.

[17] Varun Chandola, Arindam Banerjee, and Vipin Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[18] Hilmi E Egilmez and Antonio Ortega, "Spectral anomaly detection using graph-based filtering for wireless sensor networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 1085–1089.

[19] Elisabeth Drayer and Tirza Routtenberg, "Detection of false data injection attacks in smart grids based on graph signal processing," *IEEE Systems Journal*, 2019.

[20] Raksha Ramakrishna and Anna Scaglione, "Detection of false data injection attack using graph signal processing for the power grid," in *2019 IEEE Global Conference on Signal and Information Processing*, 2019.

[21] Francesco Verdoja and Marco Grangetto, "Graph laplacian for image anomaly detection," *Machine Vision and Applications*, vol. 31, no. 1, pp. 1–16, 2020.

[22] Leman Akoglu, Hanghang Tong, and Danai Koutra, "Graph based anomaly detection and description: a survey," *Data mining and knowledge discovery*, vol. 29, no. 3, pp. 626–688, 2015.

[23] Gabriela Lewenfus, Wallace Alves Martins, Symeon Chatzinotas, and Björn Ottersten, "On the use of vertex-frequency analysis for anomaly detection in graph signals," *Anais do XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2019.

[24] Yael Katz, Kolbjørn Tunstrøm, Christos C Ioannou, Cristián Huepe, and Iain D Couzin, "Inferring the structure and dynamics of interactions in schooling fish," *Proceedings of the National Academy of Sciences*, vol. 108, no. 46, pp. 18720–18725, 2011.

[25] Matthew Berger, Lee M Seversky, and Daniel S Brown, "Classifying swarm behavior via compressive subspace learning," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016.

[26] Liguo Qin, Xiao He, and DH Zhou, "A survey of fault diagnosis for swarm systems," *Systems Science & Control Engineering: An Open Access Journal*, vol. 2, no. 1, pp. 13–23, 2014.

[27] Hyojung Ahn, Han-Lim Choi, Minguk Kang, and Sung-Tae Moon, "Learning-based anomaly detection and monitoring for swarm drone flights," *Applied Sciences*, vol. 9, no. 24, pp. 5477, 2019.

[28] Yoshiki Kuramoto, "International symposium on mathematical problems in theoretical physics," *Lecture notes in Physics*, vol. 30, pp. 420, 1975.

[29] James A Hanley and Barbara J McNeil, "The meaning and use of the area under a receiver operating characteristic (roc) curve.," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

AN ENSEMBLE LEARNING FRAMEWORK FOR MULTI-CLASS COVID-19 LESION SEGMENTATION FROM CHEST CT IMAGES

Nastaran Enshaei[1], *Parnian Afshar*[1], *Shahin Heidarian*[2], *Arash Mohammadi*[1] *Moezedin Javad Rafiee*[3], *Anastasia Oikonomou*[4], *Faranak Babaki Fard*[5], *Konstantinos N. Plataniotis*[6], *and Farnoosh Naderkhani*[2],

[1]Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada
[2]Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada
[3]Department of Medicine and Diagnostic Radiology, McGill University, Montreal, QC, Canada
[4]Department of Medical Imaging, Sunnybrook Health Sciences Centre, Toronto, Canada
[5]Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[6]Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada

ABSTRACT

The novel Coronavirus disease (COVID-19) has been the most critical global challenge over the past months. Lung involvement quantification and distinguishing the types of infections from chest CT scans can assist in accurate severity assessment of COVID-19 pneumonia, efficient use of limited medical resources, and saving more lives. Nevertheless, visual assessment of chest CT images and evaluating the disease severity by radiologists are expensive and prone to error. This paper proposes an automated deep learning (DL)-based framework for multi-class segmentation of COVID lesions from chest CT images that takes the CT images as the input and generates a mask indicating the infection regions. The infection regions are segmented under two classes of data, GGOs and consolidations, which are the most common CT patterns of COVID-19 pneumonia. The proposed end-to-end framework contains four encoder-decoder-based segmentation networks that exploit the top-performing pre-trained CNNs as the encoder paths and are developed and trained separately. The results then are aggregated using a pixel-level *Soft Majority Voting* to obtain the final class membership probabilities for each pixel of the image. The proposed framework is evaluated using an open-access CT segmentation dataset. The experimental results indicate that our method successfully performs multi-class segmenting of COVID-19 lung infection regions and outperforms previous works.

Index Terms— COVID-19, segmentation, deep learning, medical imaging, ensemble-learning

1. INTRODUCTION

Over the past year, the COVID-19 pandemic has devastatingly changed many aspects of people's lives across the world. According to Johns Hopkins University (JHU) 's COVID-19 dashboard, 2,753,125 people have died due to COVID-19 pneumonia up to the 25th of March 2021 [**?**]. Hopefully, the vaccination would help end the pandemic. However, healthcare experts and authorities need to learn from this experience and be well prepared for the potential future ones.

Because of its high contingency and the increasing number of critically ill patients, severity assessment and outcome prediction of COVID-19 patients can help physicians and healthcare experts allocate limited medical resources more efficiently, make informed treatment decisions, and save more lives. Medical imaging demonstrates informative features of the COVID-19 disease and can play an essential role in pandemic management. Nevertheless, visual assessment of chest medical images and evaluating the disease severity by radiologists are expensive and prone to error. Therefore, there is an unmet need to develop automatic models for the prognosis of COVID-19 patients to speed up the severity assessment process. The most common COVID-19 manifestations in chest medical images are Ground Glass Opacities (GGOs) and consolidations. The GGO is a lung infection region with a slight increase in attenuation that does not obscure the underlying vascular system. Consolidation is a pulmonary infection with higher attenuation than GGOs in which underlying vessels and airway walls are obscured. The pure GGO is more frequently appeared in the early stages of the COVID-19 disease, whereas the observation of GGOs with consolidation is more common during progressive stages [1]. Segmenting the COVID-19 lung infection regions and detecting their types can help identify the disease severity/stage [2]. Reference [3] quantifies the percentage of consolidation areas in the whole lung volumes and introduces it as a COVID-19 severity measure.

Over the last months, many research studies have developed DL-based models for automatic diagnosis of COVID-19 patients from other Community-Acquired Pneumonia (CAP) and normal cases [4–11]. In contrast, fewer DL-based mod-

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

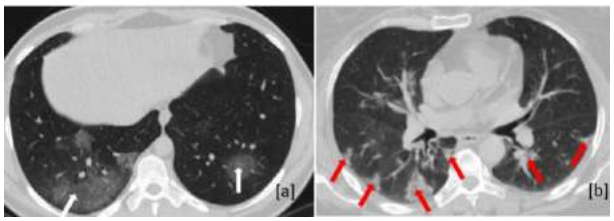


Fig. 1. The most common infection patterns in COVID-19 pneumonia in chest CT images. (a) GGOs pattern in axial CT image of a 38-year-old COVID-19 patient. (b) consolidation pattern with peripheral distribution in axial CT image of a 60-year-old woman with COVID-19 [2]

els are focusing on the segmentation of COVID-19 infection regions [12–16]. DL-based segmentation models can automatically segment the COVID-19 lung abnormalities from chest images. As the most informative chest images in representing COVID-19 lung abnormalities, CT scans have been widely used in COVID-19 lesion segmentation studies.

Segmentation models in the context of COVID-19 have been mainly developed based on U-net network [17]. U-net model, which has shown exemplary medical image segmentation results, contains an encoding path for extracting high-resolution features from images and a decoding path for learnable upsampling and reconstructing a mask indicating infection regions. Zhou *et al.* develop a U-net-based lesion segmentation model with the integration of spatial and channel attention mechanism. Reference [18] proposes a segmentation model based on U-net structure with residual connections in encoding path and a multi-scale feature integration block in bottleneck to segment COVID-19 infection regions from chest CT scans. They introduce a new loss function that helps the segmentation network learn from low-quality annotated labels. DL-based segmentation models need a large amount of annotated labels to be trained successfully. Providing large medical segmentation datasets is very expensive and time-consuming. In this regard, Fan *et al.* develop a semi-supervised learning segmentation framework that is less data-demanding and can be trained on small datasets [19]. Reference [20] synthesize COVID-19 infection regions inside healthy lung images and train their segmentation model with no labeled data. The AI-based research studies in this context are still evolving, and it needs more effort to develop automated COVID-19 diagnostic/predictive models to be reliable for clinical applications.

As mentioned previously, automated quantification of lung involvement can accelerate the disease severity assessment in COVID-19 patients. Besides, identifying different types of infection patterns from CT images can help health professionals to determine the stage/severity of COVID-19 pneumonia more accurately. Motivated by these needs, here, we propose a DL-based framework for multi-class segmentation of COVID-19 lesions from chest CT scans where the network learns to predict a class label of GGOs, consolidation, or background for each pixel of a CT image. First, four segmentation networks with state-of-the-art CNNs as the encoders are developed. The outputs of these independent segmentation networks are then combined using a pixel-level *Soft Majority Voting* approach. The proposed framework is evaluated on an open-access dataset, and the results indicate that the ensemble model can improve the overall performance of the segmentation task.

The rest of the paper is organized as follows: Section 2 represents the details of the proposed framework. Section 3 describes the CT segmentation dataset used in our experiments. The experimental setting and achieved results are demonstrated in Section 4. Finally, Section 5 discusses the results and limitations and concludes the paper.

2. PROPOSED METHOD

To develop a robust DL-based framework for the segmentation of COVID-19 lesions from chest CT scans, we adopt an ensemble learning scheme based on four independent segmentation networks. The overall pipeline of the proposed segmentation method has been shown in Fig. 2. We develop four lesion segmentation networks that are trained independently. Each segmentation network contains an encoding path for extracting high-resolution features from CT images and a decoding path for localizing the extracted features and constructing the infection masks. Adoption of state-of-the-art CNN models in encoding path can help the segmentation network learn the contextual information more accurately [**?**, 3]. Following, the CNN models incorporated as encoding path in our segmentation models are described briefly.

- **Inception-V3** [21], which is a CNN network architecture developed based on the Inception family with several improvements such as factorizing 7×7 convolutions into three 3 convolutions, label smoothing, and utilizing auxiliary classifiers for faster convergence of the network.

- **Xception** [22], which is a variant of the Inception architecture that implements depth-wise separable convolutions instead of the standard Inception modules. The number of parameters in this network is the same as Inception-v3. However, it outperforms Inception-v3 on large image classifications due to the more efficient utilization of model parameters.

- **InceptionResNet-V2** [23], which combines the Inception architecture with residual connections and through experiments indicate that residual connections can significantly speed up the training process of Inception networks.

- **DenseNet-121** [24], which is a variant of densely connected convolutional networks. In DenseNets, each

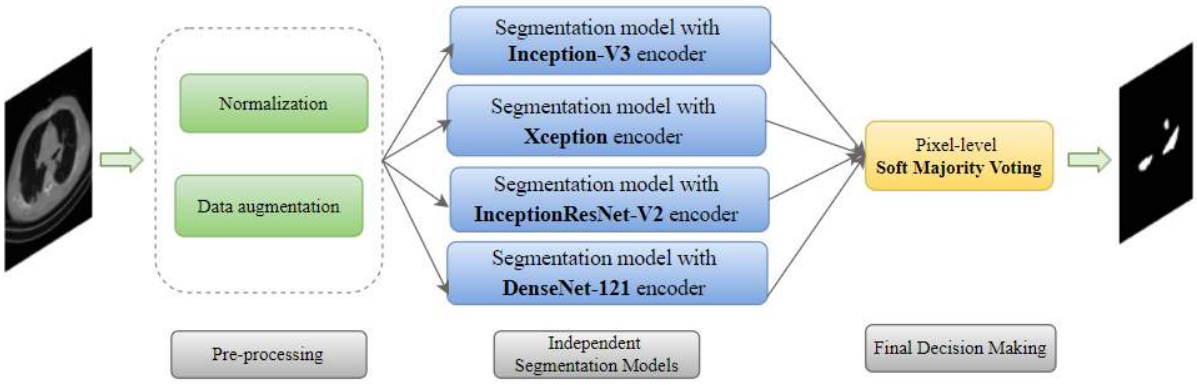


Fig. 2. The overall pipeline of the proposed COVID-19 lesion segmentation method.

layer uses the extracted feature-maps of all previous layers as the input. This dense connectivity eases the features' propagation and mitigates the vanishing-gradient problem.

We exploit the above CNN models that are available in Keras Applications alongside their pre-trained weights on the ImageNet dataset [**?**]. The fully connected layers from each pre-trained CNN model are eliminated and replaced with the decoder path to construct the segmentation networks. It is worth mentioning that any of the individual segmentation networks may outperform the others. However, since each network may have some fails and successes in predicting class labels for a class of data or a sub-group of pixels, combining them in an ensemble scheme may yield an improved performance than any of them separately. The pre-trained weights on the ImageNet are used as the network's initial weights, meaning that we keep all the layers unfrozen to train them on our dataset. The decoding path includes four decoding blocks, each consisting of a 2×2 up-sampling layer, following by a convolution layer with the kernel size of 3×3, the ReLu activation function, and the Batch-Normalization layer. ReLu activation function helps the model learn nonlinear patterns from images while avoiding the saturation of gradients. Batch-Normalization speeds up the training process by normalizing the inputs of layers and mitigating internal covariate shifts [**?**]. In the output layer, the softmax activation function predicts the probability of each pixel belongs to GGO, consolidation, or background class. The loss function is categorical cross-entropy, and the number of training epochs is 150. We use the early-stopping method to avoid over-fitting, and the training process is stopped whenever the loss function on the validation set is not decreased over ten epochs.

In the last step, we need to aggregate the outputs of four segmentation models to generate the final predicted infection masks. Voting techniques, including hard voting and soft voting, are the most commonly used when aggregating the results of multiple models. In hard voting, for each pixel of the image, the class that receives the majority of the votes from individual models is returned as the final label by the ensemble model. Soft voting includes taking the average over class membership probabilities predicted by each model where the class label with the greatest average probability is assigned as the final label. Soft voting can consider the individual models' uncertainty in ultimate decision-making by exploiting more information from predicted probabilities. Besides, pixel-level soft voting in a segmentation task (512×512 pixels for each CT image in our work) is computationally less expensive. For these reasons, we use the pixel-level soft voting method for aggregating the results of four segmentation networks as following,

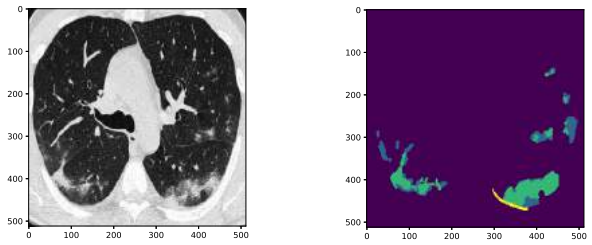
$$\hat{y} = \text{argmax}_{\text{m}} \sum_{i}^{n} w_i \, p_{i,m}, \tag{1}$$

where $\hat{y}$ is the final class label for each pixel of the image and $p_{i,m}$ is the predicted probability for class label m by the ith segmentation network. Here w_i is a weighting coefficient that determines each segmentation network's contribution in the final result. We set $w_i = \frac{1}{n}$, meaning that all segmentation models contribute equally in predicting the class labels for each pixel.

3. DATASET DESCRIPTION

The COVID-19 segmentation dataset used in our experiments is an open-access dataset provided by the Italian Society of Medical and Interventional Radiology. It includes 100 axial CT scans from 60 COVID-19 patients. The size of CT images is 512×512 pixels. A radiologist has performed the annotation of COVID-19 infection regions under three types: GGOs, consolidations, and pleural effusion. Since the regions labeled with pleural effusion contain a very small part of the images, we ignore this class of infection and limit our experiments on segmenting GGOs and consolidation infection regions, which are the most informative COVID-19 imaging manifestations. It should be noted that, to the best of our knowledge, this dataset is the only publicly available CT dataset for multi-

class segmentation of COVID-19 pneumonia. Fig 3 shows an example image of the COVID-19 segmentation dataset with its ground-truth mask.



(a) CT image (b) Ground-truth mask

Fig. 3. An example image of the COVID-19 lesion segmentation dataset with its ground-truth mask. The blue, green, and yellow areas correspond to GGOs, consolidation, and pleural effusion infection categories.

4. EXPERIMENTAL RESULTS AND DISCUSSION

We randomly split the dataset into 40, 10, and 50 images for training, validation, and testing. The images are normalized using min-max normalization. The training-testing process is performed in a 2-fold cross-validation approach. We utilize real-time data augmentation strategies, including zooming, shifting, and shearing, to avoid over-fitting. A GPU system: NVIDIA GeForce RTX 2080 Ti, CPU:i7-9700, RAM: 64 G is used in our experiments. We use the software environment: Python 3.7 Keras library, with Tensorflow backend. The model performance in segmenting GGO and consolidation infection regions is evaluated using the Dice similarity coefficient (DSC), Sensitivity (SEN), and Specificity (SPEC). The DSC metric is determined based on the coverage of predicted and ground-truth infection regions. The SEN metric measures the proportion of infection pixels correctly labeled as infection class. The SPEC metric calculates the percentage of background pixels correctly identified as background class. The evaluation metrics are defined as following, where Pr and GT correspond to the set of pixels labeled as infection regions in predicted and ground truth masks. TP, FN, TN, and FP are the number of pixels in the true positive, false negative, true negative, and false-positive regions, respectively.

$$\text{DSC} = \frac{2(|Pr| \cap |GT|)}{|Pr| + |GT|} \tag{2}$$

$$SEN = \frac{TP}{TP + FN} \tag{3}$$

$$SPC = \frac{TN}{TN + FP} \tag{4}$$

One approach when comparing multiple models based on a set of evaluation metrics is to calculate the model overall performance (MOP) using a weighted average of the metrics where the weights are determines based on the importance of each metric in the task under study. The MOP is calculated as following:

$$MOP = \sum_{i}^{n} \alpha_i \, M_i, \tag{5}$$

where MOP is the model overall performance, $\alpha_i \, \epsilon \, [0, 1]$ is the importance coefficient of metric ith, and M_i is the value of metric ith. Since in our work correctly identifying the infection pixels (measured by DSC and SEN metrics) is of more importance than detecting the background pixels (measured by the SPEC metric), we consider $\alpha_i = \{0.4, 0.4, 0.2\}$ for DSC, SEN, and SPEC, respectively. Measuring the model overall performance helps have a fair judgment when comparing multiple models' results using multiple evaluation metrics.

Table 1 presents the experimental results of each individual segmentation network and the ensemble results after *Soft Majority Voting*. The evaluation in each class of infection, GGO and consolidation, is performed separately, and then we calculate the average results. As it can be seen, the overall performance of the Inception-V3-Seg model on segmenting GGO infection regions is higher than other individual segmentation models. Simultaneously, the InceptionResNet-V2-Seg model achieves the best overall performance on the consolidation class of infection. The results indicate that the pixel-level ensemble decision-making process can help obtain higher overall performance than individual segmentation networks. We also compare our model's performance with the Reference [19] that contains two segmentation networks referred to as "Semi-Inf-Net & FCN8s" and "Semi-Inf-Net & MC". Although the "Semi-Inf-Net & FCN8s" model achieves better results regarding the DSC and SEN metrics for GGO class of infections, it almost fails in segmenting consolidation regions. Therefore, the "Semi-Inf-Net & MC" model as the most successful model proposed in Reference [19] is selected as the basis of our comparisons. The experimental results indicate that our proposed segmentation framework outperforms the "Semi-Inf-Net & MC" regarding GGO and consolidation class of infections. Fig. 4 demonstrates a visual comparison of different segmentation models on a given CT image. As can be observed, the combination of individual segmentation networks' outputs improves the final predicted mask.

5. CONCLUSION

This paper proposes a multi-class segmentation framework for segmenting COVID-19 lesions under two different classes, including GGOs and consolidation, to help health experts evaluate the disease severity more accurately. This framework contains four individual segmentation networks and a pixel-level *Soft Majority Voting* scheme for combining the results and inferring the final output. The segmentation networks

Table 1. Quantitative results of segmentation models on the test set. The best two results are highlighted in red and blue.

Methods	GGO				Consolidation				Average			
	DSC	SEN	SPEC	MOP	DSC	SEN	SPEC	MOP	DSC	SEN	SPEC	MOP
Semi-Inf-Net & MC [20]	0.624	0.618	0.966	0.69	0.458	0.509	0.967	0.5802	0.541	0.564	0.967	0.635
Inception-V3-Seg	0.536	0.573	0.978	0.639	0.511	0.603	0.982	0.642	0.523	0.588	0.980	0.640
Xception-Seg	0.618	0.652	0.981	0.704	0.524	0.451	0.994	0.589	0.571	0.552	0.987	0.647
InceptionResNet-V2-Seg	0.576	0.558	0.984	0.650	0.565	0.582	0.989	0.656	0.570	0.570	0.987	0.653
DenseNet-121-Seg	0.605	0.624	0.982	0.688	0.571	0.549	0.991	0.646	0.588	0.586	0.987	0.667
ensemble-model	0.627	0.679	0.980	0.718	0.592	0.593	0.990	0.672	0.609	0.636	0.985	0.695

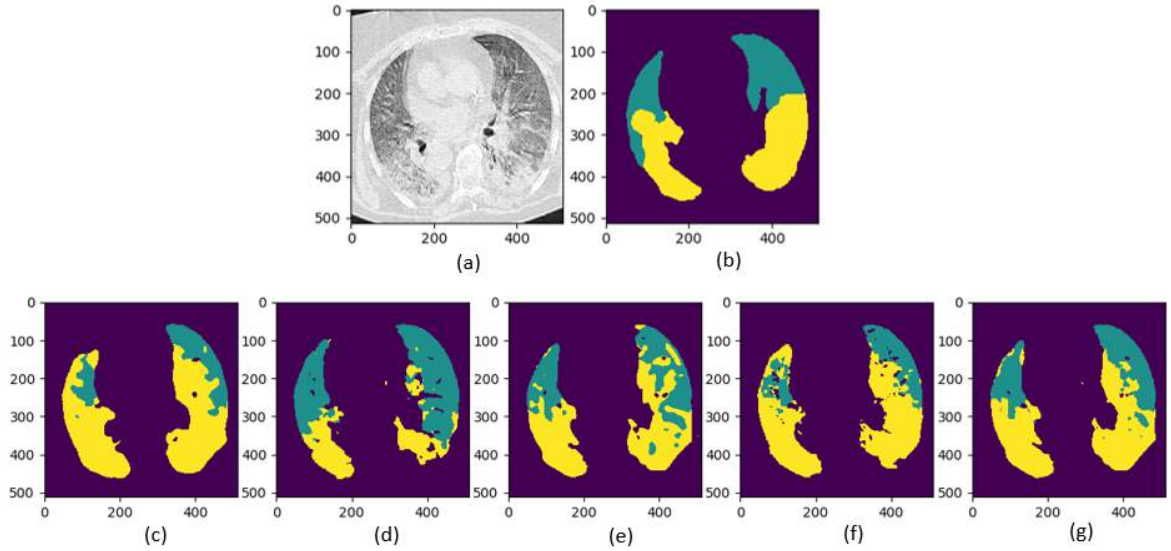


Fig. 4. Visual comparison of different segmentation models. Green and yellow represent GGOs and consolidation infection regions. (a) and (b) correspond to the original CT image and the ground truth label. (c), (d), (e), and (d) are the predicted mask by Inception-V3-Seg, Xception-Seg, InceptionResNet-V2-Seg, and DenseNet-121-Seg. (g) represents the obtained mask by the ensemble model.

are decoder-encoder-based networks that exploit state-of-the-art pre-trained CNNs, including Inception-V3, Xception, InceptionResNet-V2, and DenseNet-121 as decoder paths. The experimental results indicate that Soft Majority Voting improves the overall performance of independent segmentation networks. Our proposed framework can outperform the previous work trained and tested on the same dataset. To the best of our knowledge, the dataset we used in our experiments is the only open-access CT dataset for multi-class segmentation of COVID-19 pneumonia, which contains only 100 chest CT images alongside their infection masks. Although we tried to compensate for this problem by keeping more CT images for the test phase (50 samples were used for the training/validation set and the other 50 ones for the test set), still having a limited dataset is considered a restriction of our work.

6. REFERENCES

[1] Z. Sun, N. Zhang, Y. Li, and X. Xu, "A systematic review of chest imaging findings in covid-19," *Quantitative imaging in medicine and surgery*, vol. 10, no. 5, p. 1058, 2020.

[2] A. Mohammadi, Y. Wang, N. Enshaei, P. Afshar, F. Naderkhani, A. Oikonomou, M. J. Rafiee, H. C. Oliveira, S. Yanushkevich, and K. N. Plataniotis, "Diagnosis/prognosis of covid-19 images: Challenges, opportunities, and applications," *arXiv preprint arXiv:2012.14106*, 2020.

[3] S. Chaganti, P. Grenier, A. Balachandran, G. Chabin, S. Cohen, T. Flohr, B. Georgescu, S. Grbic, S. Liu, F. Mellot, *et al.*, "Automated quantification of ct patterns associated with covid-19 from chest ct," *Radiology: Artificial Intelligence*, vol. 2, no. 4, p. e200048, 2020.

[4] D. Singh, V. Kumar, M. Kaur, *et al.*, "Classification of covid-19 patients from chest ct images using multi-objective differential evolution–based convolutional neural networks," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 39, no. 7, pp. 1379–1389, 2020.

[5] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the covid-19 infected patients using densenet201 based deep transfer learning," *Journal of Biomolecular Structure and Dynamics*, pp. 1–8, 2020.

[6] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al Emadi, *et al.*, "Can ai help in screening viral and covid-19 pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020.

[7] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, 2021.

[8] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.

[9] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of covid-19 cases using deep neural networks with x-ray images," *Computers in biology and medicine*, vol. 121, p. 103792, 2020.

[10] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images," *Pattern Recognition Letters*, vol. 138, pp. 638–643, 2020.

[11] S. Heidarian, P. Afshar, N. Enshaei, F. Naderkhani, A. Oikonomou, S. F. Atashzar, F. B. Fard, K. Samimi, K. N. Plataniotis, A. Mohammadi, *et al.*, "Covid-fact: A fully-automated capsule network-based framework for identification of covid-19 cases from chest ct scans," *arXiv preprint arXiv:2010.16041*, 2020.

[12] Y. Qiu, Y. Liu, and J. Xu, "Miniseg: An extremely minimum network for efficient covid-19 segmentation," *arXiv preprint arXiv:2004.09750*, 2020.

[13] L. Zhou, Z. Li, J. Zhou, H. Li, Y. Chen, Y. Huang, D. Xie, L. Zhao, M. Fan, S. Hashmi, *et al.*, "A rapid, accurate and machine-agnostic segmentation and quantification method for ct-based covid-19 diagnosis," *IEEE transactions on medical imaging*, vol. 39, no. 8, pp. 2638–2652, 2020.

[14] I. Laradji, P. Rodriguez, F. Branchaud-Charron, K. Lensink, P. Atighehchian, W. Parker, D. Vazquez, and D. Nowrouzezahrai, "A weakly supervised region-based active learning method for covid-19 segmentation in ct images," *arXiv preprint arXiv:2007.07012*, 2020.

[15] Q. Yan, B. Wang, D. Gong, C. Luo, W. Zhao, J. Shen, Q. Shi, S. Jin, L. Zhang, and Z. You, "Covid-19 chest ct image segmentation–a deep convolutional neural network solution," *arXiv preprint arXiv:2004.10987*, 2020.

[16] Z. Xu, Y. Cao, C. Jin, G. Shao, X. Liu, J. Zhou, H. Shi, and J. Feng, "Gasnet: Weakly-supervised framework for covid-19 lesion segmentation," *arXiv preprint arXiv:2010.09456*, 2020.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[18] G. Wang, X. Liu, C. Li, Z. Xu, J. Ruan, H. Zhu, T. Meng, K. Li, N. Huang, and S. Zhang, "A noise-robust framework for automatic segmentation of covid-19 pneumonia lesions from ct images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2653–2663, 2020.

[19] D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "Inf-net: Automatic covid-19 lung infection segmentation from ct images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626–2637, 2020.

[20] Q. Yao, L. Xiao, P. Liu, and S. K. Zhou, "Label-free segmentation of covid-19 lesions in lung ct," *arXiv preprint arXiv:2009.06456*, 2020.

[21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[22] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, 2017.

[23] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, 2017.

[24] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

WSO-CAPS: DIAGNOSIS OF LUNG INFECTION FROM LOW AND ULTRA-LOW DOSE CT SCANS USING CAPSULE NETWORKS AND WINDOW SETTING OPTIMIZATION

Shahin Heidarian[1], Parnian Afshar[2], Nastaran Enshaei[2], Farnoosh Naderkhani[2], Moezedin Javad Rafiee, MD[3], Anastasia Oikonomou, MD[4], Faranak Babaki Fard, MD[5], Akbar Shafiee, MD[6], Konstantinos N. Plataniotis[7], and Arash Mohammadi[2]

[1]Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada
[2]Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada
[3]Department of Medicine and Diagnostic Radiology, McGill University, Montreal, QC, Canada
[4]Department of Medical Imaging, Sunnybrook Health Sciences Centre, Toronto, Canada
[5]Biomedical Sciences Department, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[6]Department of Cardiovascular Research, Tehran Heart Center,
Cardiovascular Diseases Research Institute, Tehran University of Medical Sciences, Tehran, Iran
[7]Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada

ABSTRACT

The automatic diagnosis of lung infections using chest computed tomography (CT) scans has been recently obtained remarkable significance, particularly during the COVID-19 pandemic that the early diagnosis of the disease is of utmost importance. In addition, infection diagnosis is the main building block of most automated diagnostic/prognostic frameworks. Recently, due to the devastating effects of the radiation on the body caused by the CT scan, there has been a surge in acquiring low and ultra-low-dose CT scans instead of the standard scans. Such CT scans, however, suffer from a high noise level which makes them difficult and time-consuming to interpret even by expert radiologists. In addition, some abnormalities are only visible using specific window settings on the radiologists' monitor. Currently, manual adjustment of the windowing settings is the common approach to analyze such low-quality images. In this paper, we propose an automated framework based on the Capsule Networks, referred to as the "WSO-CAPS", to detect slices demonstrating infection using low and ultra-low-dose chest CT scans. The WSO-CAPS framework is equipped with a Window Setting Optimization (WSO) mechanism to automatically identify the best window setting parameters to resemble the radiologists' efforts. The experimental results on our in-house dataset show that the WSO-CAPS enhances the capability of the Capsule Network and its counterparts to identify slices demonstrating infection. The WSO-CAPS achieves the accuracy of 92.0%, sensitivity of 90.3%, and specificity of 93.3%. We believe that the proposed WSO-CAPS has a high potential to be further utilized in future frameworks that are working with CT scans, particularly the ones which utilize an infection diagnosis step in their pipeline.

Index Terms— Low Dose CT scan, Capsule Networks, LDCT, Lung Infection

1. INTRODUCTION

The emergence of the novel coronavirus disease (COVID-19) has significantly impacted our world and made healthcare authorities facing unprecedented circumstances. Due to the highly contagious nature of the COVID-19, an early diagnosis and severity assessment of the disease will significantly help healthcare authorities to design a proper treatment plan and optimize the resources. Currently, the gold standard diagnostic tool is the Reverse Transcription Polymerase Chain Reaction (RT-PCR), which suffers from a high false negative rate [1]. This downside has led radiologists to utilize chest Computed Tomography (CT) scans as the fast and accurate alternative diagnostic tool which reveals the infection manifestations using a 3D representation of the lung constructed by a sequence of 2D images (slices) [2]. As such, CT scans have been considered as one of the primary COVID-19 diagnostic/prognostic tools in many countries. CT scans in their standard form, however, expose the patients to a high level of harmful radiation causing devastating effects on the body. Recently, Low and Ultra-Low Dose CT scans, commonly known as LDCT and ULDCT respectively, have been used in many diagnostic applications and proved to be effective in providing informative details of the disease manifestation[3, 4, 5]. Consequently, with the increasing number of suspected COVID-19 cases in need of being scanned, radiologists have introduced new protocols to acquire LDCT and ULDCT to decrease the detrimental effects of the scans caused by the radiation on the patients [6, 7]. Decreasing the radiation dose is usually performed by using lower tube currents, which in turn will impose a high level of noise on the acquired image making it difficult and time-consuming to analyze even by expert radiologists [6]. This problem has arisen the necessity of developing deep learning-based frameworks to automatically identify the disease from LDCT and ULDCT in a timely manner as well as providing additional information on the disease severity and locating the lung areas with the evidence of infection. The majority of automated frameworks using volumetric CT scans are equipped with an ROI or slice selection module in their pipeline helping them to extract slices or areas of the lung with the evidence of infection [8, 9, 10, 11]. This step plays an important role in such frameworks as it facilitates the translation from the slice-level to the patient-level domain by detecting the candidate slices or ROIs demonstrating infection at the first step, and passing them to the subsequent modules. Basically, radiologists manually adjust the screen setting using some specific windowing functions to narrow down the displayed components and adjust the image contrast as some manifestations are only visible in a specific window depending on their tissue density which is commonly distributed from $HU_{air}(-1000)$ to > 4000 in the Hounsfield

Units(HU) [12, 13]. This approach will also remove the undesired noises and artifacts in the image, facilitating its interpretation. Most windowing functions utilize mapping functions based on two parameters of Window Width (WW) and Window Level (WL) by which the function is determined.

Related Works: The majority of state-of-the-art frameworks are trained and evaluated using only standard-dose CT images and few models have been developed based on LDCT so far. As an example of such models, an end-to-end framework is proposed in [14] to predict the risk of lung cancer using 3D LDCT images. As another example, the framework developed in [15] utilizes CNN and gradient boosting decision trees to predict the risk of lung cancer, achieving the accuracy of 78.2% using LDCT images. However, no specific measure is considered in these studies to deal with the noisy and low-quality LDCT images. Recently, some research studies have incorporated a window setting optimization mechanism into the automated diagnostic/prognostic frameworks to improve the performance of the model [16, 17]. More specifically, the method proposed in [16] utilizes a stochastic window tissue normalization mechanism that randomly samples window parameters (WL,WW) from two Gaussian distributions in the training phase to segment abdominal CT images. This method, however, does not consider an optimized setting and merely normalizes the windows using randomly sampled (WL,WW). In another study [17], the proposed model uses a stack of four CNN followed by two fully connected layers as the Window Estimator Module (WEM) along with an Inception-ResNet-v2 model [13] as the lesion classifier to detect the best window setting parameter for each 2D input image. It then considers the average of obtained (WL,WW) values from the entire dataset as the final setting. Their proposed method using a combination of several window settings could improve the accuracy of the multi-class intracranial hemorrhage detection from 87.65% to 88.35% and the binary classification (Normal and Abnormal) from 95.59% to 96.43% using brain CT images. The WEM mechanism proposed in this study resulted in a wide distribution of (WL,WW) values calculated for each slice and an average function over the entire dataset might not be the optimized value. It is worthy of note that none of the aforementioned algorithms were developed based on the LDCT and ULDCT scans. In a recent study [18], a Window Setting Optimization (WSO) mechanism is proposed which uses a single convolution layer at the beginning of the pipeline to map the full-range DICOM images to the range of interest using specific windowing functions.

Contribution: In this paper, we adapted the WSO mechanism introduced in [18], to detect slices demonstrating infection in an in-house dataset of LDCT and ULDCT acquired from COVID-19 and normal cases, as well as simulated low dose images of CAP cases. More specifically, we proposed a multi-window framework, referred to as the WSO-CAPS, which applies a windowing function similar to those used by radiologist's monitors on the full-range DICOM images and passes the modified images to a classifier based on the Capsule Networks [19]. A Capsule Network-based framework has a substantial capability in capturing spatial relations between components of an image which is crucial to recognize specific patterns of the infection manifestation from medical images compared to its counterparts (e.g. CNN) [10, 20]. Using this mechanism, the WSO-CAPS identifies optimized (WL,WW) pairs that are best suited for the detection of slices demonstrating infection from LDCT and ULDCT images. We also demonstrated that it is possible to improve the performance of the classifier using an ensemble architecture to train multiple WSO modules in parallel and obtain several windowing settings at the same time. The experimental results demonstrate that the WSO-CAPS outperforms the capsule network by improving the binary (normal/abnormal) accuracy from 89.4% to 92.0%, sensitivity from 85.4% to 90.3%, and specificity from 92.2% to 93.3%. The superiority of the WSO-CAPS is also demonstrated when it is working with standard dose CT scans.

2. DATASET

In this study, we used an in-house dataset of COVID-19, CAP, and normal cases for training and evaluation of the WSO-CAPS framework. In the following subsections, the three subsets of the dataset along with a brief description of the acquisition protocols and annotation process by three experienced thoracic radiologists are described.

2.1. Low-Dose and Ultra-Low CT scans

We have collected 100 COVID-19 and 60 normal volumetric CT scans obtained from a SIEMENS SOMATOM Scope scanner in the axial view which are reconstructed using the Filtered Back Projection method [21]. The radiation dose in standard chest CT scans is estimated at $7mSv$, which is reduced to 1–$1.5mSv$ in LDCT scans and as low as $0.3mSv$ in the ULDCT ones. LDCT images are acquired from subjects with $> 60kg$ bodyweight using the mAs value of 20, kVp of $110v$, and the slice thickness of $2mm$, while the ULDCT images are obtained from subjects with the bodyweight of less than $60kg$, and the $15mAs$ has been used for the scan. This subset contains $7,703$ slices demonstrating infection and $15,464$ slices without the evidence of infection. The labeling process was performed by three experienced thoracic radiologists and the majority voting was adopted to determine the final label. This dataset is used to train, validate and evaluate the model.

2.2. Simulated Low Dose CT scans

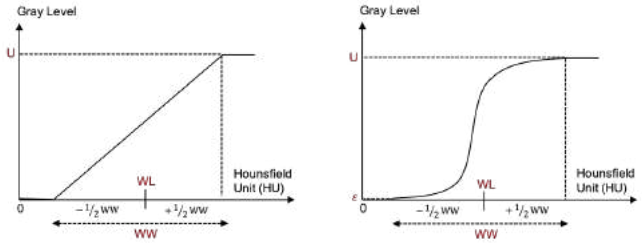
Since we did not have access to low-dose CAP CT scans, we decided to simulate them using the standard dose ones. Most of the image simulation techniques are based on paired images, which in our case means paired standard and low dose images that exactly correspond to each other. As collecting paired CT scans is not feasible for the problem at hand, we adopted an unsupervised image-to-image translation technique, referred to as CycleGan [22]. This model, essentially, consists of two sets of generators and discriminators, where the first set converts standard-dose images to low-dose ones. Consequently, the second set transfers the generated low dose images back to standard dose ones, and the output is compared with the original source image, forming the main term of the loss function. Using this technique and taking the output of the first generator, we converted standard dose CAP images to low dose ones. The dataset contains 60 simulated low dose CAP cases and is used along with the COVID-19 and normal ones to train and evaluate the model. This dataset contains $3,359$ slices demonstrating infection and $5,768$ slices without evidence of infection.

2.3. Standard Dose CT scan

To further investigate the performance of the WSO-CAPS on the standard dose CT scans, we used our in-house dataset of standard-dose CT scans referred to as the "COVID-CT-MD" [20]. The subset of the COVID-CT-MD dataset which contains slice-level labels is used as an extra test set in this study to evaluate the generalizability of the WSO-CAPS. This subset contains the slice-level labels for 54 COVID-19, 25 CAP, and 60 normal cases. We have also provided slice-level labels for 35 more CAP cases in the COVID-CT-MD dataset to expand the test set. Finally, this dataset includes $7,138$ slices demonstrating infection and $21,442$ slices without evidence of infection.

3. METHOD

In this section, the main idea behind the proposed WSO-CAPS which is the Window Setting Optimization mechanism is explained



(a) Linear Windowing Function (b) Sigmoid Windowing Function

Fig. 1. Different Windowing Functions, Figure from [13]

in detail followed by the description of the WSO-CAPS pipeline including the U-Net based lung segmentation and the Capsule Networks which are the building blocks of the WSO-CAPS framework.

3.1. Window Setting Optimization

We adopted the windowing function similar to the ones incorporated in radiologist's monitors to restrict the pixel values in a specific window ranging from 0 to the upper bound U based on the setting parameters (WL,WW). As shown in Fig. 1, linear and sigmoid mappings can be utilized as the windowing function to map all the values inside the window specified by the (WL,WW) to the [0,U], and assign all the values outside the window range to 0 or U. The linear windowing function can be formulated by the Eq. 1.

$$F_{lin}(x) = min(max(Wx + b, U), 0), \tag{1}$$

where $W = \frac{U}{WW}$ and $b = -\frac{U}{WW}(WL - \frac{WW}{2})$. The sigmoid windowing function can be formulated by the Eq. 2.

$$F_{sig}(x) = \frac{U}{1 + \exp -(Wx + b)}, \tag{2}$$

where $W = \frac{2}{WW} log(\frac{U}{\epsilon} - 1)$ and $b = \frac{-2WL}{WW} log(\frac{U}{\epsilon} - 1)$. Equations 1 and 2 indicate that the windowing function can be achieved by a convolutional layer with 1×1 filter size and a stride of 1, followed by a custom activation layer which is an upper-bounded rectified linear unit (ReLU), or sigmoid function multiplied by U, for the linear or sigmoid windowing function, respectively [18]. The proposed WSO convolutional layer can be used immediately after the input layer to display full-range DICOM images in the associated window. Using this implementation facilitates finding the optimized window settings as the weight and bias of the convolutional layer can be performed jointly with the rest of the model to determine the optimized setting (WL,WW).

3.2. WSO-CAPS

The WSO convolutional layer initiates the classification pipeline by converting the input DICOM image represented in the full-range Hounsfield Unit (i.e. ranges from -1024 to > 4000) into the specific window ranges from 0 to the upper-bound U, which is 1 in this case. Three convolution channels followed by the sigmoid windowing function are used in the WSO-CAPS model. Following the literature [23], we discarded the unrelated components of the CT images by segmenting the lung regions using a pre-trained U-Net-based lung region segmentation model [24], referred to as the "U-net (R231CovidWeb)", which has been fine-tuned specifically on the COVID-19 CT images and proved to be beneficial for the infection identification task in our previous studies [10, 11]. We also down-sampled the input images from the original 512×512 size to 256×256 to reduce the complexity and memory requirements with negligible loss of information. The result will be fed into the Capsule Network-based classifier which uses a routing by agreement process introduced in [19] to form high-level capsules (group of neurons) from low-level ones. The routing by agreement process aims to capture spatial relations between different components in an image, which is of the highest importance in the case of lung infection diagnosis. We adopted an architecture similar to our previous studies in [10, 11] which demonstrated the superiority of Capsules in the diagnosis of COVID-19 from standard dose CT images. We also added two shortcut connections to the previous model as a residual connection to further assist the model to find important features. As shown in Fig. 2, The input of the WSO-CAPS is the full-range DICOM images and the corresponding lung mask generated by the U-net (R231CovidWeb) followed by the WSO convolution layer with the size of 1×3 to detect 3 pairs of (WL,WW) at the same time. The output of the WSO layer is then fed to the stack of four convolutional layers with the size of $64, 64, 128, 128$, respectively followed by 3 layers of Capsule Networks in which the amplitude of the last layer represents the probability of the input image belonging to each target class. In addition to the aforementioned layers, a batch normalization, a pooling layer, and two shortcut connections are utilized to improve the convergence speed and generalize the learned feature maps. Finally, to deal with the unbalanced training dataset, we modified the loss function to achieve a balanced loss function by penalizing the errors caused by the misclassification of infectious slices by a higher amount. More specifically, the following equation is adopted as the balanced loss function.

$$loss = \frac{N^+}{N^+ + N^-} \times loss^- + \frac{N^-}{N^+ + N^-} \times loss^+, \tag{3}$$

where N^+ represents the number of infectious slices, N^- denotes the number of non-infectious slices, $loss^+$ is the loss value associated with infectious samples, and $loss^-$ is the loss value associated with non-infectious ones.

4. EXPERIMENTAL RESULTS

The WSO-CAPS model is trained based on the CT images acquired from 154 cases(70%) of Low Dose, Ultra-Low Dose, and simulated Low Dose CT scans from which 10% is randomly selected as the validation set to determine the best model during the training phase. It is worth mentioning that the data leakage is strictly prevented between the train and test sets. A batch size of 32, a learning rate of $1e-4$, and 100 epochs were used as the training parameters. Moreover, the weight and bias of the WSO convolution layer were initiated based on the values that corresponded to the standard windowing parameters for the lung CT scans (i.e. $WL = -500, WW = 1400$). As the first experiment, we investigated the performance of the models using sigmoid and ReLU activation functions as well as the WSO-CAPS framework without using the lung segmentation model. The corresponding results are provided in Table 1. Accordingly, we selected sigmoid windowing function using lung segmentation as the best model. In the next step, to further improve the capability of the WSO-CAPS in detecting abnormality manifestations through different windows, we performed two experiments. In the first experiment, we increased the size of the WSO convolution layer to 3, while in the second experiment, we adopted an ensemble architecture by using 3 branches of the WSO-CAPS model followed by a concatenation layer to aggregate all 3 branches in the intermediate capsule layers. We also compared the performance of the WSO-CAPS with the ResNet50 model used in [23]. The related results are presented in Table 2. Tables 1 and 2 indicate that the WSO-CAPS framework with 1 branch and 3 WSO convolution channels using the sigmoid windowing function outperforms its counterparts. They also demon-

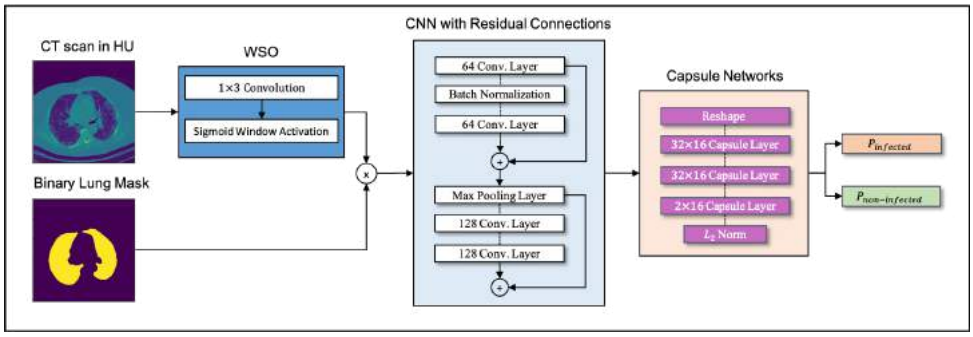


Fig. 2. WSO-CAPS Pipeline, $\times$ sign represents the element-wise multiplication, $+$ sign denotes the residual addition.

Table 1. Binary classification results obtained from the Capsuel Networks (CapsNet) and single channel WSO-CAPS.

Performance	CapsNet	CapsNet (+Residual Connection)	WSO-CAPS (ReLU)	WSO-CAPS (sigmoid)	WSO-CAPS (sigmoid, no lung segmentation)
Accuracy(%)	89.4	89.5	91.4	**91.6**	90.3
Sensitivity(%)	85.5	86.3	**91.7**	89.1	85.7
Specificity(%)	92.2	91.9	91.2	93.5	**93.9**

Table 2. Results obtained from the different architectures of the WSO-CAPS using sigmoid window activation.

Performance	WSO-CAPS (3 channels)	WSO-CAPS (3 Branches)	WSO-CAPS (3 Channels - 3 Branches)	ResNet50 (Ref [23])
Accuracy(%)	**92.0**	91.0	91.5	83.1
Sensitivity(%)	**90.3**	88.5	88.4	76.4
Specificity(%)	93.3	92.8	**93.7**	88.0

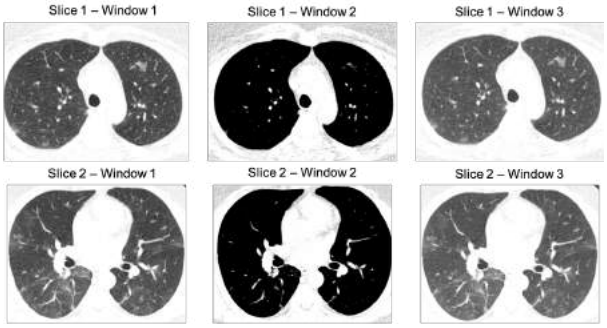


Fig. 3. Effects of three Optimized Window Settings identified by the WSO-CAPS on two sample slices.

strate that all the models equipped with the WSO mechanism perform better than the same models without using the windowing layer. It is worth mentioning that increasing the complexity of the framework by adding more convolution channels and branches could not further improve the performance. In the last step, to provide a better insight into the window setting optimization module proposed in this study, we investigated the identified (WL,WW) pairs and reviewed the CT images through the obtained window settings, considering the $\epsilon = 0.01$ in Eq. 2. The optimized setting parameters obtained by the WSO-CAPS framework using 1 channel is $(-555.9, 1032.0)$, which is quite similar to the standard setting but adds more contrast and noise reduction power to the model. The WSO-CAPS using 3 channels obtained $(-592.4, 1095.7)$, $(-277.1, 517.8)$, and $(-630.4, 1165, 4)$ as the optimized parameters. The first identified window setting in this case, is also close to the standard one which helps the model not to miss the details evident through the standard window. To better visualize the effects of the obtained parameters, Fig. 3 illustrates two sample CT images displayed by the optimized settings obtained by the WSO-CAPS using 3 channels. The capability of the WSO-CAPS to view the lung entities through different windows is evident in Fig. 3. It can also be concluded that the second window focuses more on the structure of the lung and vessels and removes the noisy and infectious components, while the first and third windows visualize the infection manifestations in different contrast levels. In another experiment, we have trained the WSO-CAPS framework using standard-dose CT scans to further investigate the generalizability of the model. In this case, the WSO-CAPS achieved the accuracy of 91.6%, sensitivity of 92.0%, and specificity of 91.4% while the CapsNet model showed a lower performance by achieving the accuracy of 90.5%, sensitivity of 89.8%, and specificity of 90.7%. Therefore, similar to the Low Dose CT scans, the superiority of the WSO-CAPS over its counterparts is evident when dealing with standard-dose CT scans.

5. CONCLUSION

In this paper, we proposed an automated framework based on the Capsule Networks, referred to as the "WSO-CAPS" to identify slices demonstrating infection from low and ultra-low-dose volumetric CT scans. The WSO-CAPS framework benefits from a Window Setting Optimization (WSO) module which is implemented by a 1×3 convolution layer followed by a sigmoid-based window activation function. The experimental results on an in-house dataset indicate that incorporation of the WSO module into the classification models will improve the performance. The proposed WSO-CAPS in this paper improved the accuracy of the Capsule Network-based classifier by 2.6%, and achieved the accuracy of 92.0%, the sensitivity of 90.3%, and the specificity of 93.3%. We also showed that the WSO-CASP using 3 WSO convolution channels will provide better results compared to using a single channel. We would like to mention that as detecting infectious slices in a volumetric CT scan is an integral step in many state-of-the-art models working with CT scans to limit the process on a small subset of candidate slices or ROIs, the WSO-CAPS will have a high potential to be incorporated in other models to improve their overall performance.

6. REFERENCES

[1] T. Ai, "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, aug 2020.

[2] M. Carotti, F. Salaffi, P. Sarzi-Puttini, A. Agostini, A. Borgheresi, D. Minorati, M. Galli, D. Marotto, and A. Giovagnoni, "Chest CT features of coronavirus disease 2019 (COVID-19) pneumonia: key points for radiologists," *La radiologia medica*, vol. 125, no. 7, pp. 636–646, jul 2020.

[3] C.M. Dorneles, G.S. Pacini, M. Zanon, S. Altmayer, G. Watte, M.C. Barros, E. Marchiori, M. Baldisserotto, and B. Hochhegger, "Ultra-low-dose chest computed tomography without anesthesia in the assessment of pediatric pulmonary diseases," *Jornal de Pediatria*, vol. 96, no. 1, pp. 92–99, jan 2020.

[4] M. Messerli, T. Kluckert, M. Knitel, S. Wälti, L. Desbiolles, F. Rengier, R. Warschkow, R.W. Bauer, H. Alkadhi, S. Leschka, and S. Wildermuth, "Ultralow dose CT for pulmonary nodule detection with chest x-ray equivalent dose – a prospective intra-individual comparative study," *European Radiology*, vol. 27, no. 8, pp. 3290–3299, aug 2017.

[5] L.J.M. Kroft, L. van der Velden, I.H. Girón, J.J.H. Roelofs, A. de Roos, and J. Geleijns, "Added Value of Ultra–low-dose Computed Tomography, Dose Equivalent to Chest X-Ray Radiography, for Diagnosing Chest Pathology," *Journal of Thoracic Imaging*, vol. 34, no. 3, pp. 179–186, may 2019.

[6] S.M.H. Tabatabaei, H. Talari, A. Gholamrezanezhad, B. Farhood, H. Rahimi, R. Razzaghi, N. Mehri, and H. Rajebi, "A low-dose chest CT protocol for the diagnosis of COVID-19 pneumonia: a prospective study," *Emergency Radiology*, vol. 27, no. 6, pp. 607–615, dec 2020.

[7] S. Tofighi, S. Najafi, S.K. Johnston, and A. Gholamrezanezhad, "Low-dose CT in COVID-19 outbreak: radiation safety, image wisely, and image gently pledge," *Emergency Radiology*, vol. 27, no. 6, pp. 601–605, dec 2020.

[8] O. Gozes, M. Frid-Adar, N. Sagie, H. Zhang, W. Ji, and H. Greenspan, "Coronavirus Detection and Analysis on Chest CT with Deep Learning," *arXiv*, apr 2020.

[9] T. Javaheri, M. Homayounfar, Z. Amoozgar, R. Reiazi, F. Homayounieh, E. Abbas, A. Laali, A.R. Radmard, M.H. Gharib, S.A.J. Mousavi, O. Ghaemi, R. Babaei, H.K. Mobin, M. Hosseinzadeh, R. Jahanban-Esfahlan, K. Seidi, M.K. Kalra, G. Zhang, L.T. Chitkushev, B. Haibe-Kains, R. Malekzadeh, and R. Rawassizadeh, "CovidCTNet:an open-source deep learning approach to diagnose covid-19 using small cohort of CT images," *npj Digital Medicine*, vol. 4, no. 1, pp. 29, dec 2021.

[10] S. Heidarian, P. Afshar, N. Enshaei, F. Naderkhani, M.J. Rafiee, F. Babaki Fard, K. Samimi, S.F. Atashzar, A. Oikonomou, K.N. Plataniotis, and A. Mohammadi, "COVID-FACT: A Fully-Automated Capsule Network-Based Framework for Identification of COVID-19 Cases from Chest CT Scans," *Frontiers in Artificial Intelligence*, vol. 4, may 2021.

[11] S. Heidarian, P. Afshar, A. Mohammadi, M.J. Rafiee, A. Oikonomou, K.N. Plataniotis, and F. Naderkhani, "Ct-Caps: Feature Extraction-Based Automated Framework for Covid-19 Disease Identification From Chest Ct Scans Using Capsule Networks," in *ICASSP 2021 IEEE International Conference on Acoustics, Speech and Signal Processing*. jun 2021, pp. 1040–1044, IEEE.

[12] K.T. Bae, G.N. Mody, D.M. Balfe, S. Bhalla, D.S. Gierada, F.R. Gutierrez, C.O. Menias, P.K. Woodard, J. M. Goo, and C.F. Hildebolt, "CT Depiction of Pulmonary Emboli: Display Window Settings," *Radiology*, vol. 236, no. 2, pp. 677–684, aug 2005.

[13] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, Feb. 2017.

[14] D. Ardila, A.P. Kiraly, S. Bharadwaj, B. Choi, J.J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D.P. Naidich, and S. Shetty, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Medicine*, vol. 25, no. 6, pp. 954–961, jun 2019.

[15] J.L. Causey, Y. Guan, W. Dong, K. Walker, J.A. Qualls, F. Prior, and X. Huang, "Lung cancer screening with low-dose CT scans using a deep learning approach," *arXiv*, jun 2019.

[16] Y. Huo, Y. Tang, Y. Chen, D. Gao, S. Han, S. Bao, S. De, J.G. Terry, J.J. Carr, R.G. Abramson, and B.A. Landman, "Stochastic tissue window normalization of deep learning on computed tomography," *arXiv*, dec 2019.

[17] M. Karki, J. Cho, E. Lee, M.H. Hahm, S.Y. Yoon, M. Kim, J.Y. Ahn, J. Son, S.H. Park, K.H. Kim, and S. Park, "CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings," *Artificial Intelligence in Medicine*, vol. 106, pp. 101850, jun 2020.

[18] H. Lee, M. Kim, and S. Do, "Practical Window Setting Optimization for Medical Image Deep Learning," *arXiv*, dec 2018.

[19] G. Hinton, S. Sabour, and N. Frosst, "Matrix capsules with EM routing," *6th International Conference on Learning Representations, ICLR 2018*, pp. 1–29, 2018.

[20] P. Afshar, S. Heidarian, N. Enshaei, F. Naderkhani, M.J. Rafiee, A. Oikonomou, F. Babaki Fard, K. Samimi, K.N. Plataniotis, and A. Mohammadi, "COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning," *Scientific Data*, vol. 8, no. 1, pp. 121, dec 2021.

[21] F. Pontana, J. Pagniez, T. Flohr, J. Faivre, A. Duhamel, J. Remy, and M. Remy-Jardin, "Chest computed tomography using iterative reconstruction vs filtered back projection (Part 1): evaluation of image noise reduction in 32 patients," *European Radiology*, vol. 21, no. 3, pp. 627–635, mar 2011.

[22] J.Y. Zhu, T. Park, P. Isola, and A.A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*. oct 2017, pp. 2242–2251, IEEE.

[23] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, and J. Xia, "Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, aug 2020.

[24] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, and G. Langs, "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem," *European Radiology Experimental*, vol. 4, no. 1, pp. 50, dec 2020.

MULTI-SLICE NET: A NOVEL LIGHT WEIGHT FRAMEWORK FOR COVID-19 DIAGNOSIS

Harshala Gammulle, Tharindu Fernando, Sridha Sridharan, Simon Denman, Clinton Fookes

The Signal Processing, Artificial Intelligence and Vision Technologies (SAIVT),
Queensland University of Technology, Australia.

ABSTRACT

This paper presents a novel lightweight COVID-19 diagnosis framework using CT scans. Our system utilises a novel two-stage approach to generate robust and efficient diagnoses across heterogeneous patient level inputs. We use a powerful backbone network as a feature extractor to capture discriminative slice-level features. These features are aggregated by a lightweight network to obtain a patient level diagnosis. The aggregation network is carefully designed to have a small number of trainable parameters while also possessing sufficient capacity to generalise to diverse variations within different CT volumes and to adapt to noise introduced during the data acquisition. We achieve a significant performance increase over the baselines when benchmarked on the SPGC COVID-19 Radiomics Dataset, despite having only 2.5 million trainable parameters and requiring only 0.623 seconds on average to process a single patient's CT volume using an Nvidia-GeForce RTX 2080 GPU.

Index Terms— COVID19 Diagnosis, Deep Learning, Computed Tomography, Medical Imaging.

1. INTRODUCTION

Although Reverse Transcription Polymerase Chain Reaction (RT-PCR) is considered the global standard SARS-CoV-2 (COVID-19) diagnosis, this test is very time consuming and has a high false negative rate, which in turn yields significant challenges in preventing the spread of the infection [1, 2]. As such, Computed Tomography (CT) imaging has been identified as a fast, simple and reliable diagnosis tool due to the existence of discriminative patterns associated with the COVID-19 infection within the CT scans. However, recent literature has shown that COVID-19 lung manifestations show substantial similarities with Community Acquired Pneumonia (CAP), complicating the diagnosis process [1].

To this end several deep learning based frameworks have been introduced to automate diagnosis, where models are trained to uncover discriminative patterns embedded within the data and which cannot be identified by the naked-eye. This paper presents the QUT SAIVT team's [1] framework for the 2021 IEEE ICASSP Signal Processing Grand Challenge (SPGC) – "COVID-19 Radiomics". This challenge dataset has been constructed to motivate machine learning practitioners to develop robust and reliable systems to classify patients into COVID-19, CAP and NORMAL diagnosis classes using a heterogeneous set of CT scans. In particular, these CT scans are composed of different slice thicknesses, radiation doses, and noise levels, in addition to featuring patients with various comorbidities and different surgical histories.

While volumetric CT scans provide a comprehensive illustration of lung abnormalities and their structure, patient level diagnosis from heterogeneous CT volumes faces several challenges as noise and variation between scans can lead to misclassification of individual CT slices. Hence, simplistic score-level/ feature-level [1, 3] aggregation performs poorly as there is a tendency for some slices to be misclassified. Structures such as 3D-CNNs have also been used to regress volumetric CT inputs directly to the final diagnosis decision [2, 4, 5]. While this allows the model to extract and operate over feature vectors that represent the entire lung of the patient, these models have a very high-dimensional parameter space (tens of millions of trainable parameters) and are prone to over-fitting when trained using datasets with patients (individual samples) in the order of hundreds.

To alleviate these challenges we propose a novel two-stage framework where features from individual slices are aggregated to a patient level diagnosis via an efficient, light-weight 1D-CNN based model. As novel contributions, (1) our design exploits slice-level features from adjacent slices at different granularities, combining and compressing these discriminative features, prior to classification; (2) has fewer trainable parameters, enabling effective training from a smaller set of volumetric CT scans; (3) our method allows us to seamlessly process examples with a variable number of slices, and even allows the model to learn from incomplete/partial scans; and (4) due to the use of a pre-trained backbone (feature extractor) to extract features from the individual CT slices, the backbone can be swapped or modified. Hence, the proposed two-stage framework is not limited to CT lung classification tasks, but can be easily adapted to any diagnosis task which requires aggregation of heterogeneous information across different samples.

[1] https://research.qut.edu.au/saivt/

2. SPGC COVID-19 RADIOMICS DATASET

The SPGC COVID-19 Radiomics Dataset is one of the largest datasets containing COVID-19, Community Acquired Pneumonia (CAP), and normal cases, and is captured in different medical centers with various imaging settings. The dataset comprises volumetric CT scans of 307 patients (171 COVID-19, 60 CAP, and 76 NORMAL patients). All captured slices in the CT scans are in the Digital Imaging and Communications in Medicine (DICOM) format. The data is acquired using a SIEMENS, SOMATOM Scope scanner with the normal radiation dose and the slice thickness of 2mm. Apart from this patient level labelling, a small subset (i.e 55 COVID-19, and 25 CAP) were analyzed and the individual slices were labeled to indicate evidence of infection. In total 4,993 slices were identified as being indicative of infection. From this dataset, 30% of the data was randomly selected and provided as a validation set. The validation set contains 98 patients (55 COVID, 19 CAP, and 24 NORMAL). The test set consists of three subsets where they consist of 35 COVID, 20 CAP, and 35 NORMAL patients. Test dataset labels are withheld, however, we report the challenge evaluation released by the organisers.

3. METHODOLOGY

We propose a deep network approach, Multi-slice Net, which performs the lung infection classification from the volumetric chest CT scans. The proposed framework is shown in Fig. 1, and is composed of a backbone for slice level feature extraction and a network to aggregate these features from a patient to a single score (Multi-Slice Network).

3.1. Feature Extractor/ Backbone

One of the key motivations of the proposed approach is to minimise pre-processing. Hence, aside from converting individual DICOM files to JPG format, no pre-processing steps are performed. In contrast to existing state-of-the-art approaches [1, 2] which perform lung detection and segmentation during pre-processing, the proposed framework applies the feature extractor directly to the JPG slice images.

Extracting features from CT slices that capture discriminative infection-related information is crucial for infection classification. We utilise the squeeze-and-excitation ResNet50 (SE-ResNet50) model [6], pre-trained on the ImageNet dataset [7].The SE-ResNet50 extends the original ResNet50 architecture with the aid of squeeze and excitation operations. In particular, the squeeze operation extracts global information from each of the channels of the input while the excitation act as a bottleneck, adaptively recalibrating the importance of each channel. We fine-tune the SE-ResNet50 model, though the first 6 layers are frozen. For fine-tuning, the subset of patients with slice level annotations are used. This subset contains 55 COVID, and 25 CAP patients. We also randomly selected slices from 15 NORMAL patients for the fine-tuning data. The constructed dataset contains of 2482 COVID, 742 CAP, and 1820 NORMAL slices for training and 1333 COVID, 436 CAP and 840 NORMAL slices for validation. For the compatibility with the pre-trained backbone network, the input CT slices of shape $512 \times 512 \times 1$ are resized to $224 \times 224 \times 1$ and replicated 3 times ($224 \times 224 \times 3$), before being fed to the backbone SE-ResNet50. To reduce over-fitting we used data augmentation and added Random Horizontal Flips with 50% probability, and randomly changed the brightness, contrast and saturation of the input by a factor of upto 0.4. The network is trained using the Adam [8] optimiser with a learning rate of $1e{-}5$ using Categorical Cross-Entropy Loss for 100 epochs. We used class weights to balance the impact of the minority classes.

After fine-tuning, we use the model with best validation accuracy and extract the features from the penultimate layer of SE-ResNet50, with a feature dimensionality of 2048.

3.2. Multi-Slice Network

The features extracted from the backbone then fed to the proposed Multi-Slice Net to obtain a patient-level infection classification. Fig. 2 illustrates our approach. Multi-Slice Net iterates through the features extracted from all the CT slices that belong to a particular patient, aggregating them, and generates a single feature descriptor representing the patient. This feature is then fed through a series of dense layers to obtain the final patient diagnosis.

In the challenge dataset, the CT scans of each patient have a varying number of slices. As such, we designed the network to handle a variable number of slices with the aid of fully convolution network. Let the number of CT slices for a particular subject be l, then, the input (I) of the Multi-Slice Net takes the shape (l, 2048). Our Multi-Scale Net consists of a temporal convolution layer followed by 4 dilated residual blocks, composed of dilated convolutions. Inspired by [9, 10], we doubled the dilation factor at each layer, and the number of convolutional filters used at each layer is 64. The output of the fourth dilated residual block is then passed through a max-pooling layer with the kernel size of l, which encodes the input sequence into a single feature with dimensionality of 64. This feature is then passed through the classification network which is composed of two dense layers with sizes 32 and 3 (number of infection classes) respectively. The use of temporal convolution allows our network to interrogate the slice level features at different granularities, comparing and contrasting features of neighbouring slices. By aggregating these features to a single vector, the most salient features from the patient are passed to the classifier. This network is trained using the Adam [8] optimiser with a learning rate of $1e{-}4$ using Categorical Cross-Entropy Loss for 100 epochs.

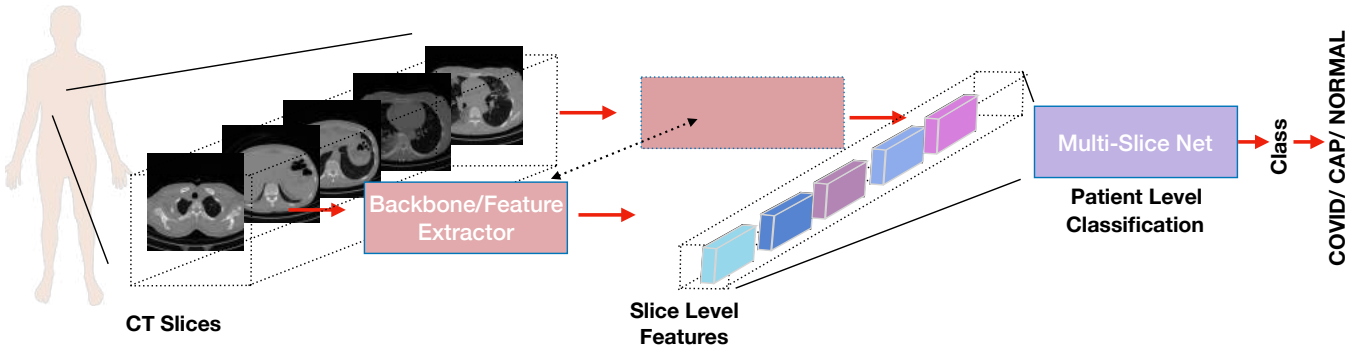


Fig. 1. Overall Framework: Individual volumetric chest CT slices are passed through a backbone network for slice level feature extraction. The resultant features are aggregated by the proposed Multi-Slice Network to obtain a patient level diagnosis.

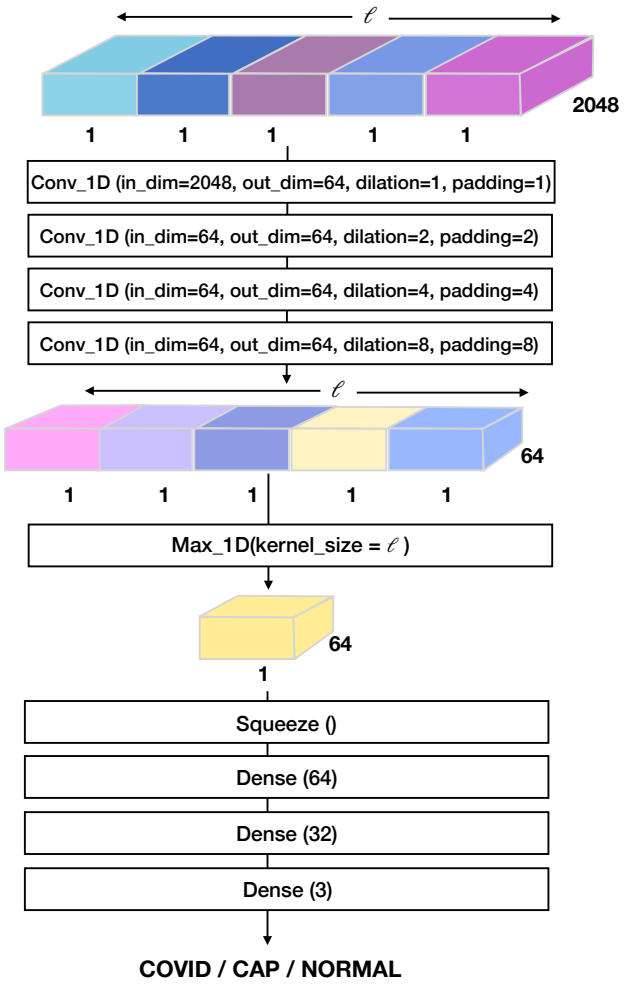


Fig. 2. Multi-Slice Network

4. EVALUATION RESULTS

In this section, we first present evaluation results for the fine-tuning process of the feature extractor (Sec. 4.1). In Sec. 4.2 we report patient level diagnosis performance using Multi-Slice Net (MS-Net).

4.1. Slice Level Classification Performance (Backbone Networks)

We evaluate several network architectures to determine an appropriate backbone for feature extraction. When fine-turning these networks, we initialised them with their respective ImageNet weights and fine-tuned them for 100 epochs using the Adam optimiser, a learning rate of $1e{-}5$ and the categorical cross entropy loss. Note that for the fine-tuning process we utilised a subset of the SPGC COVID-19 Radiomics Dataset provided by the organisers which had slice level annotations (see Sec. 3.1 for details).

Method	Validation Sensitivity			Validation Accuracy
	COVID	CAP	NORMAL	
DenseNet [11]	32.80%	83.66%	71.52%	63.52%
ResNet-18 [12]	60.84%	61.39%	90.82%	71.02%
SqueezeNet [13]	76.72%	71.06%	89.95%	79.15%
ResNet-50 [12]	72.08%	78.47%	88.80%	79.92%
SE-ResNet-50 [6]	**79.50%**	**84.58%**	**96.02%**	**86.63%**

Table 1. Slice-level classification accuracy using CT slices from a subset of the SPGC COVID-19 Radiomics dataset. We report class-level sensitivity for COVID, Community Acquired Pneumonia (CAP), and NORMAL classes and overall accuracy (percentage of correct predictions).

Tab. 1 provides results for the ResNet-18 [12], ResNet-50 [12], SqueezeNet [13], DenseNet [11] , and SE-ResNet-50 [6] architectures when fine-tuned to obtain a slice level diagnosis. We observe superior performance from the SE-ResNet-50 architecture, despite of the fact that it has been introduced for channel level feature re-calibration on RGB inputs. Despite the need to replicate a single channel CT slice image three times to satisfy the 3-channel requirement of the network, we observe a significant performance increase between ResNet-50 and SE-ResNet-50. We believe this is a result of the removal of redundant/replicated information in

channels through the squeeze and excitation blocks of SE-ResNet-50, allowing the classification layers to better focus on informative spatial attributes of the input.

4.2. Patient-Level Evaluation

Evaluation results with respect to the validation set of SPGC COVID-19 Radiomics dataset are provided in Tab. 2. We report the results of the baseline model provided by the challenge organisers as well as results for MS-Net with different backbones. Our framework outperforms the baseline system, especially when considering the COVID detection sensitivity. We observe similar performance between the ResNet and SE-ResNet backbones, despite the significant performance gap between these methods with respect to slice level evaluations. In Tab. 3 we provide results across testing subsets of the SPGC COVID-19 Radiomics dataset. Despite the lightweight architecture we observe that our framework has achieved competitive results for all classes across all subsets. As the ground truth labels of the test data is not available we cannot compare our performance with existing state-of-the-art models. However, we note that this framework achieved 9th place (from 17 competitive systems) in the SPGC COVID-19 Radiomics challenge. Furthermore, one important characteristic of the proposed method is its consistent performance across the different classes. Despite the heterogeneous test sets, including different slice thicknesses, radiation dose, patient level differences, our lightweight system has been able to achieve consistent performance.

Method	Validation Sensitivity			Validation Accuracy
	COVID	CAP	NORMAL	
Baseline (Provided by Challenge Organisers)	42.10 %	**94.5%**	75.00%	79.60%
MS-Net with ResNet-50 backbone	**87.76%**	86.67%	77.27%	**84.88%**
MS-Net with SE-ResNet-50 backbone	**87.76%**	66.67%	**90.91%**	**84.88%**

Table 2. Patient-level evaluations using the validation set of the SPGC COVID-19 Radiomics dataset. We report class-level sensitivity scores for the COVID, Community Acquired Pneumonia (CAP), and NORMAL classes as well as the overall accuracy in terms of the percentage of correctly predicted observations.

Test subset	COVID	CAP	NORMAL	Total
Test set 1	13/15	NA	14/15	27/30
Test set 2	4/10	10/10	5/10	19/30
Test set 3	7/10	9/10	10/10	26/30

Table 3. Patient-level evaluations on different test subsets of the SPGC COVID-19 Radiomics dataset. We report the number of correct identifications against the total ground truth examples for each class. NA refers to Not Applicable as no examples were present in that particular subset.

Another noteworthy aspects of the proposed approach is the ability to seamlessly switch between different backbone networks. Due to our two-stage approach, the architecture of MS-Net does not require any changes when changing the backbone feature extractor. Moreover, the backbone can be trained in a separate dataset, even without any patient-level data (i.e multiple-slices per patient). As the proposed MS-Net has fewer trainable parameters it can be tuned later with a small scale dataset with fewer patient-level annotations. In addition, we highlight that the MS-Net architecture is not limited to slice level feature aggregation from CT scans. It could be utilised for any aggregation task where features from different spatial or temporal locations need to be aggregated.

4.3. Network Complexity

The majority of the trainable parameters in our framework lie within the backbone feature extractor (SE-ResNet-50), which has 2.5 million trainable parameters (the first six layers are frozen during fine-tuning). MS-Net has only 207,683 trainable parameters due to its careful design. Despite the parameter heavy design of the backbone, the plug and play nature of MS-Net allows the backbone to be pre-trained on a completely different data corpus, and fine-tuned for the task at hand using a smaller dataset. It generates 268 patient level predictions (each of which has a variable number of slices, between 100 and 200, per patient) in 166.9619 seconds. This includes inference for both the backbone network for feature extraction and MS-Net to obtain patient-level predictions. Therefore, on average it takes only 0.6229 seconds to process a CT volume. In future works we will be investigating better backbone architectures to further improve our model's performance, while maintaining it's light weight nature.

5. CONCLUSION

We present a novel light weight framework for COVID-19 diagnosis. Our approach uses a two-stage architecture, composed of a backbone network for feature extraction from individual CT scan slices, and a network to aggregate these slice level features for patient-level diagnosis. Considering the limited data availability of complete patient-level CT volumes, we design a light-weight network to aggregate the slice-level features for patient-level diagnosis. This system is evaluated using the SPGC COVID-19 dataset and achieves competitive results. One prominent attribute of our design is the plug-and-play nature of the aggregation network, which allows the backbone to be trained on a completely different dataset and then tuned on a smaller dataset for the task at hand with patient-level annotations. Future work will include investigation of other backbone designs to further improve model accuracy while maintaining its light weight nature.

6. REFERENCES

[1] Shahin Heidarian, Parnian Afshar, Arash Mohammadi, Moezedin Javad Rafiee, Anastasia Oikonomou,

Konstantinos N Plataniotis, and Farnoosh Naderkhani, "Ct-caps: Feature extraction-based automated framework for covid-19 disease identification from chest ct scans using capsule networks," *arXiv preprint arXiv:2010.16043*, 2020.

[2] Chuansheng Zheng, Xianbo Deng, Qing Fu, Qiang Zhou, Jiapei Feng, Hui Ma, Wenyu Liu, and Xinggang Wang, "Deep learning-based detection for covid-19 from chest ct using weak label," *MedRxiv*, 2020.

[3] Arash Mohammadi, Yingxu Wang, Nastaran Enshaei, Parnian Afshar, Farnoosh Naderkhani, Anastasia Oikonomou, Moezedin Javad Rafiee, Helder CR Oliveira, Svetlana Yanushkevich, and Konstantinos N Plataniotis, "Diagnosis/prognosis of covid-19 images: Challenges, opportunities, and applications," *arXiv preprint arXiv:2012.14106*, 2020.

[4] Hasib Zunair, Aimon Rahman, Nabeel Mohammed, and Joseph Paul Cohen, "Uniformizing techniques to process ct scans with 3d cnns for tuberculosis prediction," in *International Workshop on PRedictive Intelligence In MEdicine*. Springer, 2020, pp. 156–168.

[5] Huseyin Polat and Homay Danaei Mehr, "Classification of pulmonary ct images by using hybrid 3d-deep convolutional neural network architecture," *Applied Sciences*, vol. 9, no. 5, pp. 940, 2019.

[6] J Hu, L Shen, S Albanie, G Sun, and E Wu, "Squeeze-and-excitation networks.," *IEEE transactions on pattern analysis and machine intelligence*, 2019.

[7] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[8] Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[9] Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu, "Wavenet: A generative model for raw audio," *ISCA Speech Synthesis Workshop (SSW)*, 2016.

[10] Yazan Abu Farha and Jurgen Gall, "Ms-tcn: Multi-stage temporal convolutional network for action segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3575–3584.

[11] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[13] Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.

Using reinforcement learning to forecast the spread of COVID-19 in France

Soheyl Khalilpourazari
Department of Mechanical, Industrial, and Aerospace Engineering
Concordia University
Montreal, Canada H3G 1M8
soheyl.khalilpourazari@mail.concordia.ca

Hossein Hashemi Doulabi
Department of Mechanical, Industrial, and Aerospace Engineering
Concordia University
Montreal, Canada H3G 1M8
hossein.hashemi@concordia.ca

***Abstract*— In December 2020, a new strain of coronavirus was found in Wuhan, China. The virus causes COVID-19, a severe respiratory illness. Up to date, the virus has spread rapidly to many countries, and more than 103 million cases and 2 million death has been reported worldwide. France is one of the European Union countries that has reported more than 3 million cases and 76 thousand death. Prediction of the COVID-19 pandemic growth is essential to enable governments to put new measures to slow down the spread of the virus. Due to the virus's novelty, providing an efficient method to predict pandemic growth is a challenging task. This research applies a recent reinforcement learning-based algorithm to a recently developed model to simulate the COVID-19 pandemic in France. We provide essential information about the pandemic growth in the country in every period in which the government of France has taken action to limit the pandemic or relaxed existing restrictions. We derive the values of the pandemic parameters, including reproduction rate, which gives us essential information about the pandemic. This information will help policymakers and healthcare professionals to plan for future measures limiting community transmission. Besides, we performed sensitivity analyses to determine the most critical parameters that accelerate the pandemic.**

Keywords—Reinforcement learning, SIDARTHE, pandemic modeling, covid-19, machine learning.

I. INTRODUCTION

A new strain of coronavirus was discovered in Wuhan, China, in December 2020. The new virus causes COVID-19, which is a severe respiratory illness. Despite initial restrictions put in place by the government, the virus spread to several countries worldwide. Up to date, more than 103 million cases and 2 million death has been reported worldwide. To limit the pandemic growth and limit community transmission, it is crucial to predict the pandemic growth by estimating some critical epidemiological parameters, which is a challenging task due to the novelty of the virus and the limited number of researches done to simulate the COVID-19 pandemic.

Over the last year, many researchers aimed at proposing efficient methods to model and simulate pandemic growth. Some researchers applied pure machine learning methods to predict the COVID-19 pandemic. Zhang et al. [1] presented a Poisson formulation to predict the upcoming number of daily cases. Using the new methodology, researchers estimated the peak time of the pandemic. Chimmula and Zhang [2] used a Long-Short Term Memory (LSTM) based methodology to predict the COVID-19 pandemic. Arora et al. [3] also used an LSTM to simulate the COVID-19 pandemic in India. Ogundokun and Awotunde [4] applied Support Vector Regression, Neural Network, and Linear Regression to predict the COVID-19 pandemic in India. The authors declared that these methodologies perform well in predicting the upcoming peak of the pandemic. Tuli et al. [5] proposed a new machine learning-based algorithm using a repetitive weighting following Generalized Inverse Weibull distribution to predict the COVID-19 pandemic. The authors performed the simulation on cloud computing for better results. Malki et al. [6] used the regressor machine learning method to predict the pandemic growth and approximate pandemic parameters. The authors' main novelty was considering the environmental conditions on virus transmission and determining the effect of these factors on pandemic growth. Many other researchers used machine learning to predict the COVID-19 pandemic [7-12]. Although these methods perform well in fitting a function to real data to predict the pandemic, they have some limitations. For instance, these methods cannot provide the decision-makers with epidemiological information of the pandemic, such as transmission rate and reproduction rate.

Some other researchers used differential equations to predict and model the COVID-19 pandemic. These methods have several advantages compared to machine learning methods. For instance, they can provide the decision-makers with important epidemiological information. These methods can accurately estimate the pandemic parameters. One of the most famous pandemics modeling models is called the SIR model, which is widely used in the literature to model outbreaks of several diseases [13]. The SIR model only considers susceptible, recovered, and infected cases and has several disadvantages. To overcome these limitations and provide an efficient mathematical expression of the COVID-19 pandemic, Giordano et al. [14] enhanced the SIR model by considering other essential factors and developed the SIDARTHE model. The SIDARTHE considers susceptible (S), infected (I), diagnosed (D), ailing (A), recognized (R), threatened (T), healed (H), and extinct (E) cases. The authors implemented the model to real data from Italy and showed that the model is able to model the COVID-19 pandemic very efficiently. This model has one limitation, and it is the computational burden of the solution process. The solution of the SIDARTHE model is a challenging task that limits its application in reality. Khalilpourazari and Hashemi Doulabi [15] proposed a Hybrid Q-Learning-based Algorithm to predict the COVID-19 pandemic in Quebec, Canada. They showed

978-1-7281-7289-7/20/$31.00 ©2020 IEEE

that their algorithm performs very well in forecasting pandemic growth.

The current research applies a new Reinforcement Learning (RL) based procedure to resolve the SIDARTHE model in a reasonable time. By hybridizing the SIDARTHE model and machine learning, we benefit from both methods and provide a practical methodology to predict and model the COVID-19 pandemic. We apply our methodology to real data from France to predict the COVID-19 pandemic. We obtain the values of the epidemiological parameters and provide a detailed guide on how tightening and relaxing social measures affect the pandemic growth in the country. We provide detailed information about the reproduction rate to depict the pandemic growth. In the end, we perform sensitivity analyses to study the effect of changes in the pandemic variables growth. These data will help the policymakers and healthcare professionals in planning for further restrictions to fight pandemic growth.

II. Solution Methodology

Based on the No-Free Lunch (NFL) theorem, there is no single metaheuristic that performs the best in all optimization problems [16]. This is due to the fact that randomization plays a prominent role in the performance of metaheuristics. The efficiency of a metaheuristic in solving a given problem meaningfully depends on the problem's solution space and the optimization framework of the algorithm. The theory behind updating a particle in the solution space is the most crucial factor in the performance of these algorithms. Some algorithms use direct movement, such as Particle Swarm Optimization (PSO), and some may use encircling strategies such as Moth-Flame Optimization (MFO). Each optimization paradigm has its advantages and limitations. However, finding a way to switch between several operators of different algorithms intelligently would result in a higher ability to search the solution space.

Khalilpourazari and Hashemi Doulabi [15] proposed a new algorithm called the Hybrid Q-Learning-based Algorithm (HQLA) that uses several metaheuristics to solve the most complex benchmarks. The authors used a reinforcement learning-based method called Q-learning that acts as the main algorithm in the suggested method. RL is a subset of machine learning that refers to algorithms that interact with the environment to maximize a reward [17]. These algorithms aim to find an optimal strategy to minimize punishment and maximize reward.

Q-Learning is an efficient RL algorithm used to achieve the best policy to maximize the expected reward. This algorithm aims to determine the best state-action pairs using (1).

$$Q_{(t+1)}(s_t,a_t)=Q_t(s_t,a_t)+\mathcal{E}_t(r_t+\gamma\max(Q_t(s_{t+1},a_{t+1}))-Q_t(s_t,a_t)), \quad (1)$$

where r_t is the amount of reward/punishment, γ is a scaling factor, and $\mathcal{E}_t$ is the learning rate [18]. This algorithm repeats over the iterations and updates the Q-table to choose the best action in the current state. Algorithm 1 shows the pseudo-code of the Q-learning.

Algorithm 1: Q-Learning

```
input;
primary state;
For the number of iterations do
        determine the action;
        perform the action
        calculate reward/punishment;
        chose action;
        update the table;
        Go to the new state;
  end
  return;
```

The HQLA uses several different updating paradigms (actions) to perform optimization. Using the Q- learning algorithm, it chooses the best action when updating the position of the particles in the solution space during the optimization process. Throughout iterations, the algorithm learns to optimize operator selection to optimize its performance. Algorithm 2 shows the pseudo-code of the HQLA.

In Algorithm 2, the GWO, SFS, WCA, PSO, MFO, and SCA stand for Grey Wolf Optimizer, Stochastic fractal Search, Water Cycle Algorithm, Particle Swarm Optimization, Moth-Flame Optimization, and Sine Cosine Algorithm, respectively [19-25].

Algorithm 2: HQLA [15]

```
generate an initial solution set;
For the number of iterations
   correct infeasibility;
   calculate the objective value (error);
   for 1:npop
      if iteration =1
         Random action;
      else
         Chose an action using Q-learning;
      end
      if action =1
        apply GWO;
      else if action =2
        apply SFS;
      else if action =3
        apply WCA;
      else if action =4
        apply PSO;
      else if action =5
        apply MFO;
      else if action =5
        apply SCA;
      end
   correct infeasibility;
   calculate the objective value (error);
   calculate reward/punishment;
   find the max Q;
   update a;
   update Qt(st,at);
   t=t+1;
   end
end
return;
```

III. Case Study

France is part of the ongoing worldwide COVID-19 pandemic caused by the new virus discovered in late 2019. The virus, which is called SARS-Cov-2, is a new strain of coronavirus which causes severe respiratory illness. Due to the high transmissibility of the virus, it spread rapidly to many countries. France reported 3,177,879 infected cases and 75,862 deaths by January 31, 2021. The first case of COVID-19 was reported in France on January 24, 2020. On March 12, the government of France announced that the country would go into a limited-time lockdown to limit the spread of the virus. Since then, the government has placed several lockdown periods to limit the spread of the virus. In December 2020, the UK announced a new and more infectious and transmissible variant of the COVID-19. On December 25, 2020, France confirmed the 1st case of the new variant (B117) of the COVID-19. Due to the high transmissibility of the new variant of COVID-19, France faced a surge in daily new cases. Consequently, many individuals were admitted to hospitals and Intensive Care Units (ICUs). This showed the importance of new and efficient methodologies to predict the pandemic growth and take action before facing a new surge in the number of new cases.

To take action and plan for future lockdowns, governments need to predict some parameters of the pandemic, such as the reproduction rate. For this purpose, Giordano et al. [14] proposed a new mathematical model to simulate the pandemic growth of the COVID-19. The model is called SIDARTHE, which models many pandemic parameters and gives valuable information to policymakers, healthcare professionals, and immunologists. Although the model is an efficient mathematical expression of pandemic growth, solving such complex differential equations for many iterations (periods, e.g., days) is challenging. To solve the model and apply it to real data from France, we used a novel algorithm called HQLA. We used recent data on the COVID-19 pandemic in France, including the number of daily new cases, recovered cases, and death cases. We note that these data are publicly available on https://data.humdata.org/event/covid-19.

To estimate the outbreak parameters, we first considered the data from January 22, 2020 (day 1) to January 21, 2021 (day 366). Then we break this period into several sub-periods in which the government of France applied or relaxed restrictions and lockdowns. Then we used the HQLA algorithm to fit the predictions to real data using the mathematical model of the SIDARTHE.

It is disclosed that the HQLA achieves a proper trade-off between exploration and exploitation of the solution space. This helps the HQLA avoid trapping in local optima and achieve a high-quality solution for the problem with a mean square error of 2.25E-04. Based on the outcome of the HQLA, we estimated the pandemic parameters from day 1 to day 366, as provided in Tables I and II.

TABLE I. Results of the SIDARTHE and HQLA.

Parameters	Stages				
	January 22, to March 13	March 13 to March 30	March 30, to May 11	May 11 to August 1	August 1 to October 5
α	0.170726	0.031766	0.004165	0.004165	0.059832
β	0.331386	0.020619	0.010710	0.010710	0.019706
δ	0.003431	0.020619	0.010710	0.010710	0.019706
γ	0.331386	0.027542	0.001387	0.001387	0.058781
ε	0.001308	0.001308	0.000354	0.000354	0.000354
ζ	0.004489	0.004489	0.004489	0.004417	0.004417
η	0.004489	0.004489	0.004489	0.004417	0.004417
θ	0.003532	0.003532	0.003532	0.003532	0.003532
λ	0.003106	0.003106	0.003106	0.003106	0.009823
κ	0.086303	0.086303	0.086303	4.40E-05	4.40E-05
ξ	0.086303	0.086303	0.086303	4.40E-05	4.40E-05
ρ	0.003106	0.003106	0.003106	4.40E-05	4.40E-05
σ	0.086303	0.086303	0.086303	4.40E-05	4.40E-05
μ	0.02251	0.02251	0.02251	0.000479	0.000479
υ	0.047352	0.047352	0.047352	0.047352	0.000243

TABLE II. RESULTS OF THE SIDARTHE AND HQLA (CONTINUED).

Parameters	Stages				
	October 5 to November 5	November 5 to November 12	November 12 to November 28	November 28 to December 15	After December 15
α	0.113324	0.010956	0.000227	0.00478	0.029501
β	0.028410	0.008967	0.000986	0.006625	0.019778
δ	0.028410	0.008967	0.000986	0.006625	0.019778
γ	0.007158	0.003709	2.10E-05	0.000554	0.002495
ε	0.000989	0.000989	0.000989	0.000989	0.000989
ζ	2.10E-05	2.10E-05	2.10E-05	2.10E-05	2.10E-05
η	2.10E-05	2.10E-05	2.10E-05	2.10E-05	2.10E-05
θ	0.003532	0.003532	0.003532	0.003532	0.003532
λ	0.009823	0.009823	0.009823	0.009823	0.009823
κ	0.000645	0.000645	0.000645	0.000645	0.000645
ξ	0.000645	0.000645	0.000645	0.000645	0.000645
ρ	0.000645	0.000645	0.000645	0.000645	0.000645
σ	0.000509	0.000509	0.000509	0.000509	0.000509
μ	0.000479	0.000479	0.000479	0.000479	0.000479
υ	0.000243	0.000243	0.000243	0.000243	0.000243

Based on the results, we calculated the reproduction rate of the pandemic during each stage of the horizon. Our results show a significant reproduction rate in the first stage of the pandemic, equal to 20.73. Starting March 13, the government of France ordered the closure of the restaurants and other entertainment centers, which significantly dropped the reproduction rate to 4.10. On March 30, the government announced that they would further extend the lockdown to stop pandemic growth. This resulted in an even more decrease in the reproduction rate to 0.597. On May 11, the government of France announced that some schools could be reopened. The reopening caused the reproduction rate to increase to 11.17 and accelerate the pandemic growth by reopening borders to incoming flights and allowing businesses to reopen the virus started to spread more rapidly than ever, increasing the reproduction rate to 28.32 after August 1. On October 5, the government ordered the entertainment centers' closure, resulting in a decrease in the reproduction rate to 14.495. More information about our results is given in Fig. 1 to Fig. 7.

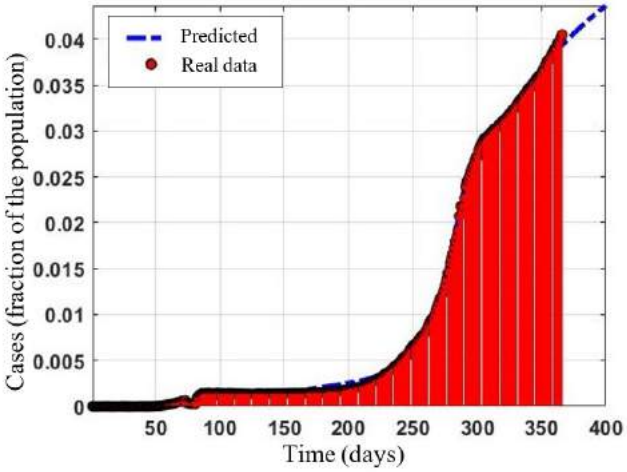


Fig. 1. The number of infected cases model vs. data.

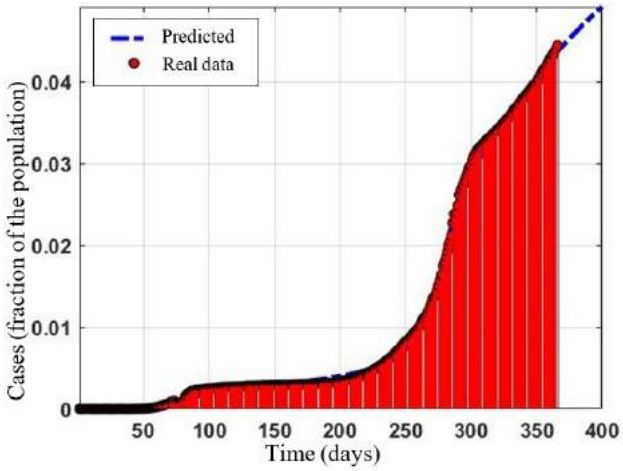


Fig. 2. The number of diagnosed cases model vs. data.

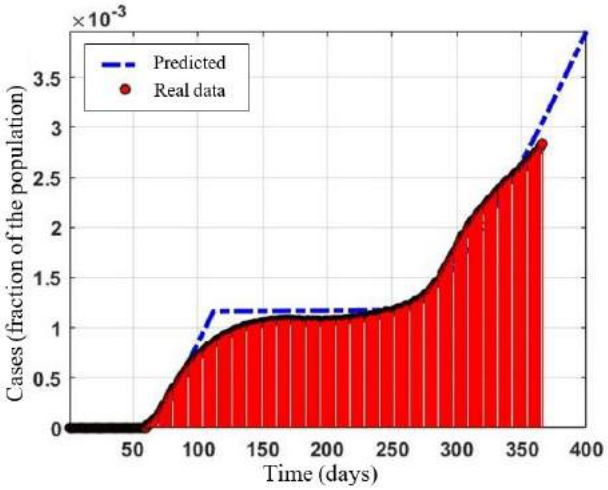


Fig. 3. The number of recovered cases model vs. data.

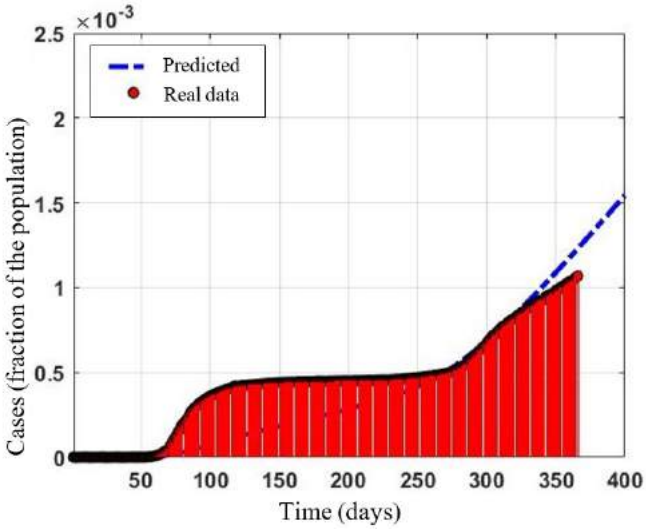


Fig. 4. The number of death cases model vs. data.

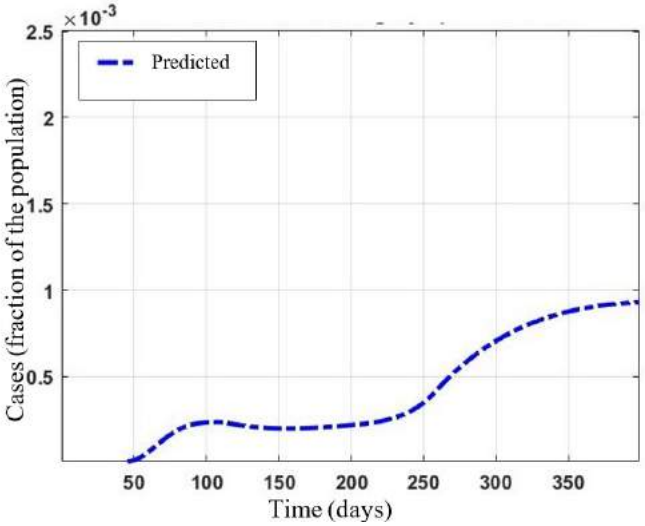


Fig. 5. The number of infected cases which will develop life-threatening symptoms model.

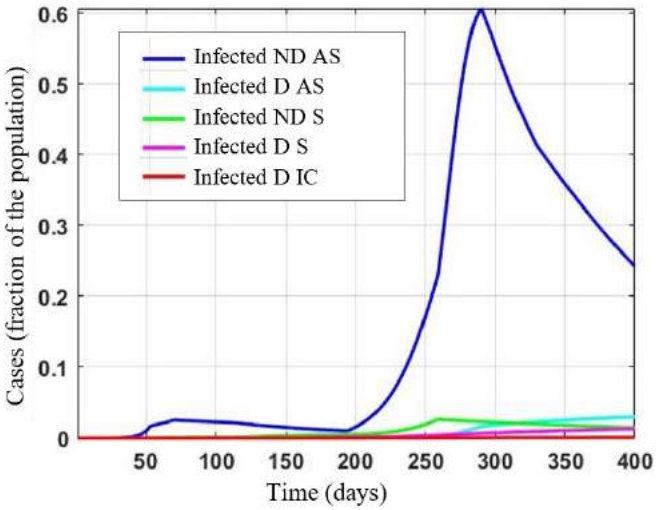


Fig. 6. Prediction of the future cases. Non-Diagnosed Asymptomatic (ND AS), Diagnosed Asymptomatic (D AS), Non-Diagnosed Symptomatic (ND S), Diagnosed Symptomatic. (D S), and Diagnosed with Life-Threatening Symp. (D IC).

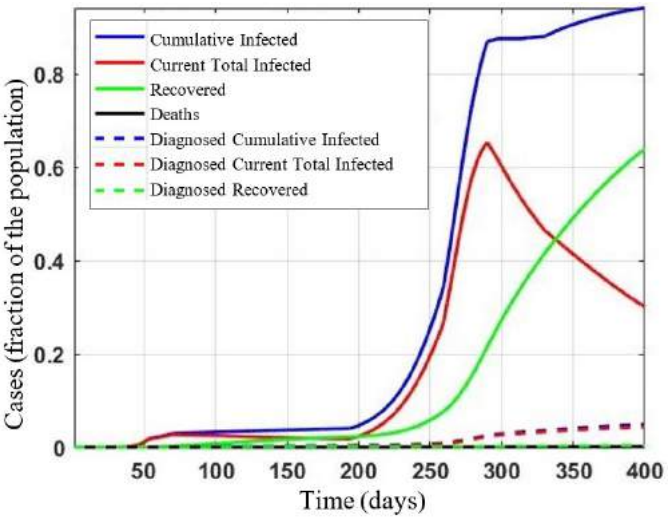


Fig. 7. Prediction of the COVID-19 pandemic in France.

On November 5 Paris Mayor declared additional limitations in Paris. This limited the reproduction rate to 1.571. This decrease continued to November 28 where the reproduction rate reached 0.076. After that, some restrictions were relaxed and caused the reproduction rate to jump to 0.834. On December 15, some travel restrictions were relaxed by the government of France and enhanced the pandemic growth resulting in a reproduction rate of 3.985. If the current restrictions remain in place, we predict an increase in the infected cases in France in the upcoming months.

IV. SENSITIVITY ANALYSES

This section evaluates the effect of any change in the transmission rate parameters on pandemic growth. This is important because we need to project the effect of new lockdowns or relaxing lockdowns on future pandemic growth. Fig.8 to 16 show the outcomes of the sensitivity analyses. The results show that the parameter Alfa has the most significant effect on pandemic growth. Therefore, it is essential to place new measures to limit community transmission and pandemic growth.

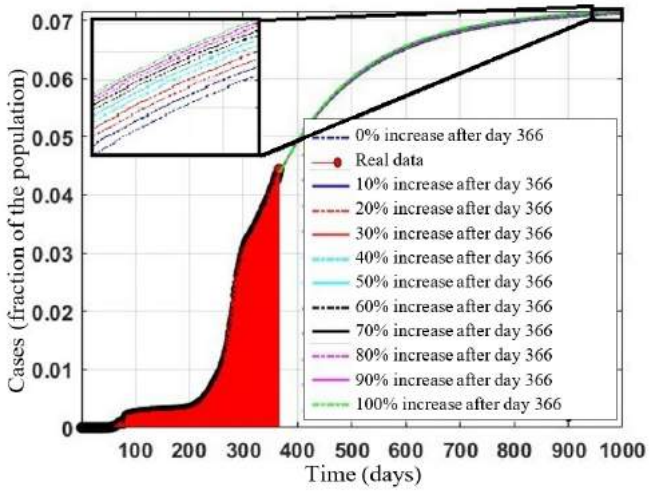


Fig. 8. Effect of change in Alfa on diagnosed cases.

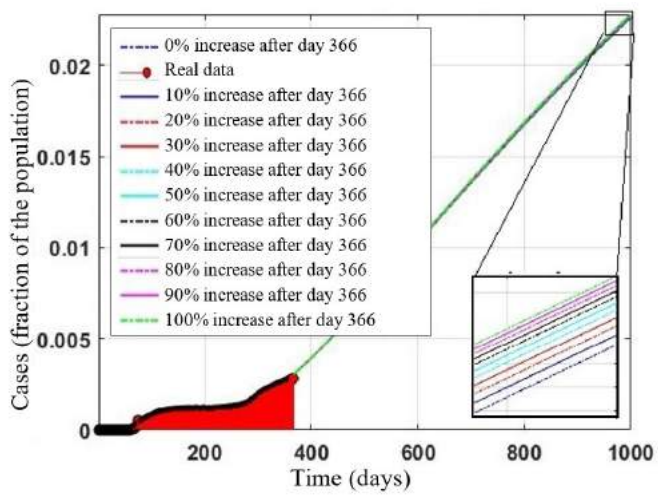


Fig. 9. Effect of change in Alfa on recovered cases.

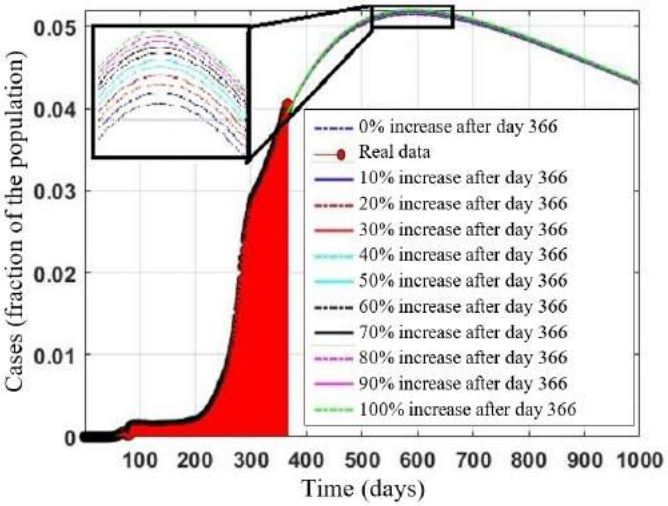


Fig. 10. Effect of change in Alfa on infected cases.

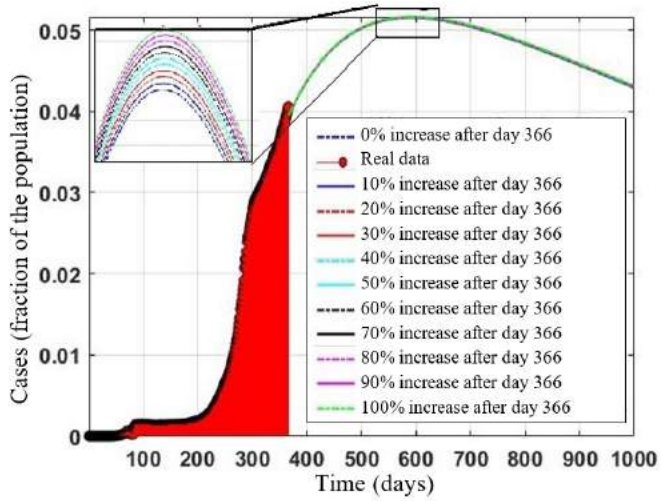


Fig. 11. Effect of change in Beta and Delta on infected cases.

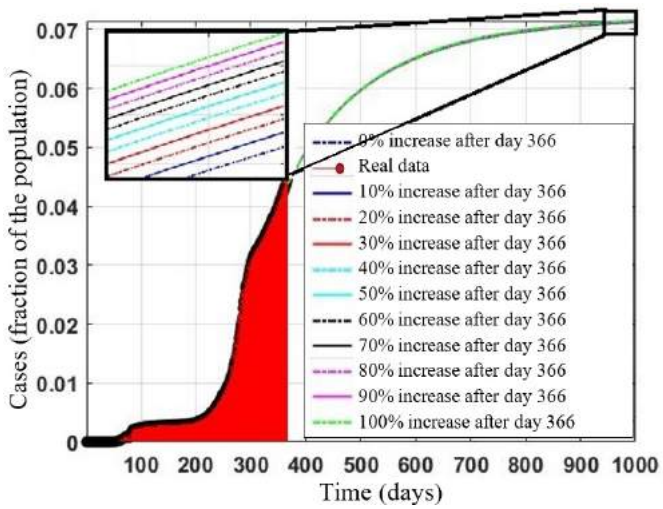


Fig. 12. Effect of change in Beta and Delta on diagnosed cases.

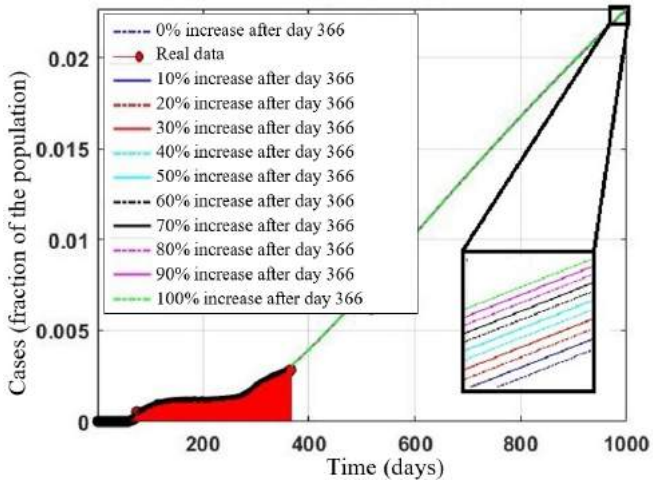


Fig. 13. Effect of change in Beta and Delta on recovered cases.

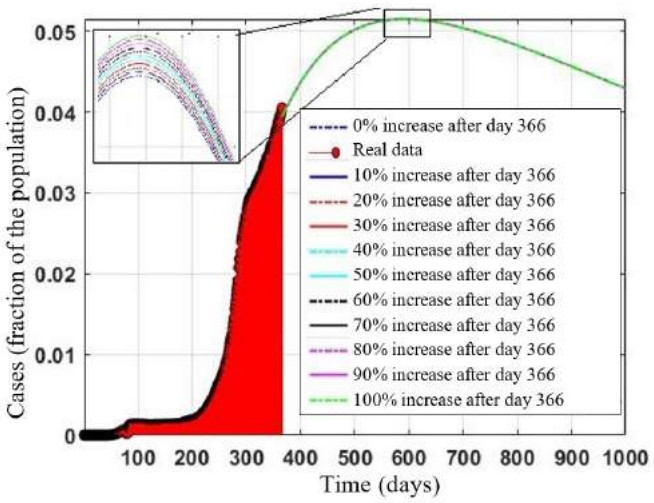


Fig. 14. Effect of change in Gamma on infected cases.

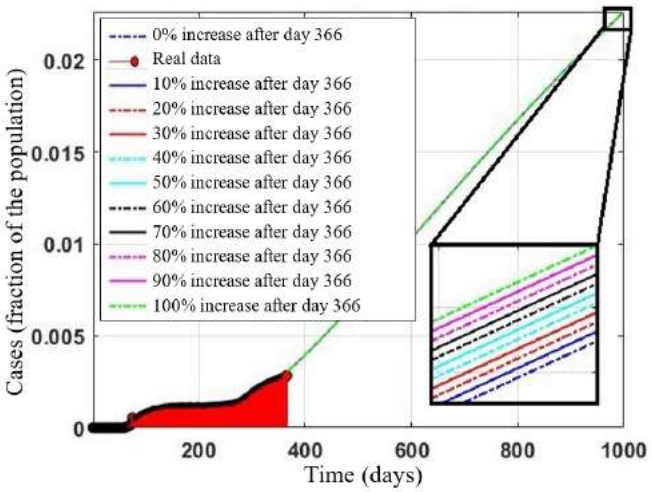


Fig. 15. Effect of change in Gamma on recovered cases.

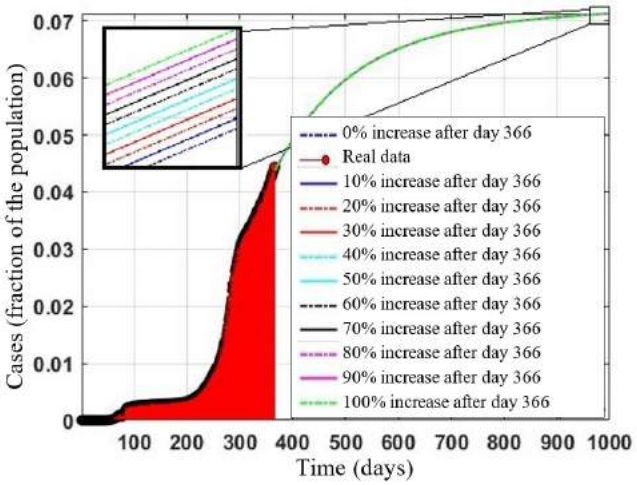


Fig. 16. Effect of change in Gamma on diagnosed cases.

REFERENCES

[1] Zhang X, Ma R, Wang L. Predicting turning point, duration and attack rate of COVID-19 outbreaks in major Western countries. Chaos, Solitons & Fractals. 2020 June 1;135:109829.

[2] Chimmula VK, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals. 2020 June 1;135:109864.

[3] Arora P, Kumar H, Panigrahi BK. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos, Solitons & Fractals. 2020 October 1;139:110017.

[4] Ogundokun RO, Awotunde JB. Machine learning prediction for covid 19 pandemic in india. medRxiv. 2020 Jan 1.Tuli S, Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. Internet of Things. 2020 Sep 1;11:100222.

[5] Tuli S, Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. Internet of Things. 2020 Sep 1;11:100222.

[6] Malki Z, Atlam ES, Hassanien AE, Dagnew G, Elhosseini MA, Gad I. Association between weather data and COVID-19 pandemic predicting mortality rate: Machine learning approaches. Chaos, Solitons & Fractals. 2020 Sep 1;138:110137.

[7] Abebe TH. Forecasting the Number of Coronavirus (COVID-19) Cases in Ethiopia Using Exponential Smoothing Times Series Model. medRxiv. 2020 January 1.

[8] Alamo T, Reina DG, Millán P. Data-driven methods to monitor, model, forecast and control covid-19 pandemic: Leveraging data science, epidemiology and control theory. arXiv preprint arXiv:2006.01731. 2020 June 1.

[9] Garcia LP, Goncalves AV, de Andrade MP, Pedebos LA, Vidor AC, Zaina R, de Luca Canto G, de Araujo GM, Amaral FV. Estimating underdiagnosis of covid-19 with nowcasting and machine learning: Experience from brazil. medRxiv. 2020 January 1.

[10] da Silva RG, Ribeiro MH, Mariani VC, dos Santos Coelho L. Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. Chaos, Solitons & Fractals. 2020 October 1;139:110027.

[11] Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos, Solitons & Fractals. 2020 Jun 25:110059.

[12] Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Singh V. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. Chaos, Solitons & Fractals. 2020 Sep 1;138:109944.

[13] Weiss HH. The SIR model and the foundations of public health. Materials matematics. 2013:0001-17.

[14] Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, Colaneri M. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nature medicine. 2020 Jun;26(6):855-60.

[15] Khalilpourazari S, Doulabi HH. Designing a hybrid reinforcement learning based algorithm with application in prediction of the COVID-19 pandemic in Quebec. Annals of Operations Research. 2021 Jan 3:1-45.

[16] Ho YC, Pepyne DL. Simple explanation of the no-free-lunch theorem and its implications. Journal of optimization theory and applications. 2002 Dec;115(3):549-70.

[17] Bertsekas DP. Reinforcement learning and optimal control. Belmont, MA: Athena Scientific; 2019 March 13.

[18] Zamli KZ, Din F, Ahmed BS, Bures M. A hybrid Q-learning sine-cosine-based strategy for addressing the combinatorial test suite minimization problem. PloS one. 2018 May 17;13(5):e0195675.

[19] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Advances in engineering software. 2014 Mar 1;69:46-61.

[20] Salimi H. Stochastic fractal search: a powerful metaheuristic algorithm. Knowledge-Based Systems. 2015 Feb 1;75:1-8.

[21] Khalilpourazari S, Naderi B, Khalilpourazary S. Multi-objective stochastic fractal search: A powerful algorithm for solving complex multi-objective optimization problems. Soft Computing. 2020 Feb;24(4):3037-66.

[22] Eskandar H, Sadollah A, Bahreininejad A, Hamdi M. Water cycle algorithm–A novel metaheuristic optimization method for solving constrained engineering optimization problems. Computers & Structures. 2012 Nov 1;110:151-66.

[23] Kennedy J, Eberhart R. Particle swarm optimization. InProceedings of ICNN'95-international conference on neural networks 1995 November 27 (Vol. 4, pp. 1942-1948). IEEE.

[24] Mirjalili S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-based systems. 2015 Nov 1;89:228-49.

[25] Mirjalili S. SCA: a sine cosine algorithm for solving optimization problems. Knowledge-based systems. 2016 Mar 15;96:120-33.

On Future Development of Autonomous Systems: A Report of the Plenary Panel at IEEE ICAS'21

Yingxu Wang, *Fellow, IEEE,* University of Calgary, AB, Canada (yingxu@ucalgary.ca)
Ioannis Pitas, *Fellow, IEEE,* Aristotle University of Thessaloniki (AUTH), Greece (pitas@csd.auth.gr)
Konstantinos N. Plataniotis, *Fellow, IEEE,* University of Toronto, ON, Canada (kostas@ece.utoronto.ca)
Carlo S. Regazzoni, *SM, IEEE,* University of Genova, Italy (carlo.regazzoni@unige.it)
Brian M. Sadler, *Life Fellow, IEEE,* The US Army Research Laboratory, USA (brian.m.sadler6.civ@mail.mil)
Amit Roy-Chowdhury, *Fellow, IEEE,* University of California, Riverside, CA, USA (amitrc@ece.ucr.edu)
Ming Hou, *SM, IEEE,* Defence Research and Development Canada, Toronto, Canada (ming.hou@drdc-rddc.gc.ca)
Arash Mohammadi, *SM, IEEE,* Univ. of Concordia, Montreal, QC, Canada (arash.mohammadi@concordia.ca)
Lucio Marcenaro, *SM, IEEE,* University of Genova, Italy (lucio.marcenaro@unige.it)
Farokh Atashzar, *MIEEE,* New York University, NY, USA (sfa7@nyu.edu)
Saif alZahir, *SM IEEE,* Univ. of Concordia, Montreal, QC, Canada (saifz@encs.concordia.ca)

Abstract – Autonomous Systems (AS) are perceived as the most advanced intelligent systems evolved from those of reflexive, imperative, and adaptive intelligence. A plenary panel on "Future Development of Autonomous Systems" is organized at the inaugural IEEE ICAS'21. This paper reports the panel discussions about the-state-of-the-art and paradigms of AS, the basic research on theoretical foundations and mathematical means of AS, and challenges to the future development of AS. As an emerging and increasingly demanded field, AS provide an unprecedented approach to contemporary intelligent industries including deep machine learning, highly intelligent robotics, cognitive computers, general AI technologies, and industrial applications enabled by transdisciplinary advances in intelligence science, system science, brain science, cognitive science, robotics, computational intelligence, and intelligent mathematics.

Keywords – *Autonomous systems, intelligence systems, general AI systems, cognitive systems, theoretical foundations, machine learning, challenges, constraints, applications*

1. Introduction

It is recognized that Autonomous Systems (AS) are advanced intelligent systems and general AI technologies triggered by the transdisciplinary development in intelligence science, system science, brain science, cognitive science, robotics, computational intelligence, and intelligent mathematics [1-5]. As an emerging field, AS address the challenges to general AI and the next generation of intelligent systems where both state spaces of their stimuli and behavioral generations are dynamically indeterministic at design time or pending for run-time. These fundamental constraints to current computing and AI theories and platforms indicate the needs not only for novel technical developments, but also for deep basic research on theoretical foundations, advanced theories, intelligent computing platforms/languages, and underpinning intelligent mathematics for rigorous modeling the unprecedented demands for cognitive computing and general AI [6-8].

The primary purpose of this plenary panel on *The Future Development of Autonomous Systems*, at IEEE 1st International conference on Autonomous Systems (ICAS'21) [1], is to provide participants for an interactive means, particularly in the environment of virtual conference, to learn from distinguished experts' perspectives towards the essences and trends in AS development. It also allows participants to obtain professional visions, insights, and feedbacks to strategic questions or fundamental challenges to AS.

This paper presents a summary of the plenary panel. The distinguished panelists represent a group of the world's preeminent scholars and experts in basic research and industrial innovations on AS. The talks, discussions, and interactions with audience show the panelists' visions and perspectives on the trend to future development of AS in the fast emerging and transdisciplinary field across intelligent science, computational intelligence, general AI, computer science, system science, intelligent mathematics, as well as engineering demands from a wide spectrum of modern industries.

2. Autonomous Systems: Basic Research and the Future Development

AS are perceived as a run-time deterministic intelligent system that depends not only on current stimuli or demands, but also on internal goals, status, and knowledge formed by historical learning and current rational inferences. The ultimate goal of AS is to implement a brain-inspired system that may think and act as a human counterpart in hybrid intelligent systems and general AI implementations. AS enable nondeterministic behaviors at run-time closer to that of humans at the level of cognitive intelligence [2, 8]. Well known and potential paradigms of AS may encompass brain-inspired AI systems such as those of deep machine learning systems, machine consciousness and awareness implementations, cognitive robots, surgical robots, self-driving vehicles,

autonomous drones, real-time machine inference engines, brain-machine interfaces, and knowledge-based intelligent systems.

A *Hierarchical Intelligence Model* (HIM) is introduced to reveal the levels of intelligence and their increasing complexities and difficulties for implementation in intelligence science and computational intelligence as shown in Figure 1 [2]. According to HIM, the levels of human and AS intelligence are aggregated from reflexive, imperative, adaptive, to autonomous and cognitive intelligence. HIM indicates that both human and machine intelligence are formed layer-by-layer from the bottom up. Without the underpinning layers, the upper layers may not be implemented. The HIM model is logically and neurologically consistent to the discovery of the *Layered Reference Model of the Brain* (LRMB) [9, 10] where the brain encompasses the following layers of natural intelligence: 1) Sensory, 2) Action, 3) Memory, 4) Perception, 5) Cognition, 6) Inference, and 7) Autonomous intelligence [32-36]. The LRMB model provides a brain/cognitive science foundation for modeling brain-inspired systems (BIS). Based on this perspective, any AS is equivalent to a BIS, or vice versa, which is essentially characterized by run-time derived intelligent behaviors beyond those of pretrained or predetermined ones at designed time.

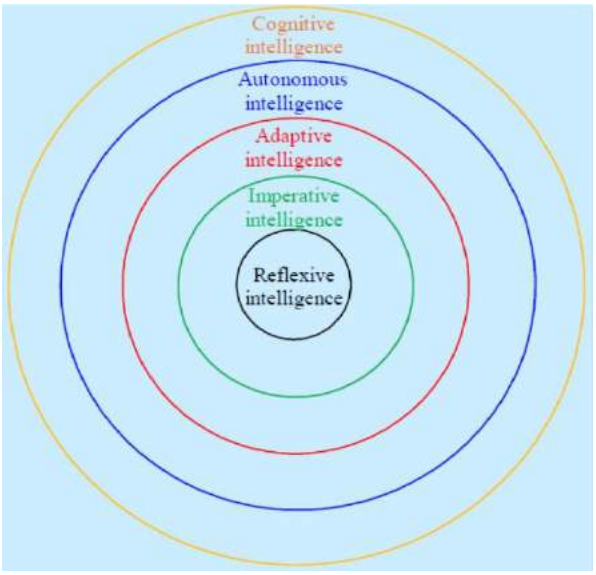


Fig. 1. The hierarchical intelligence model (HIM) of AS

In order to explore the essences and challenges to the design and implementation of AS, as well as for stimulating deep thoughts for the panel discussions, the following set of *Key Fundamental Challenges* (KFC) to AS is proposed to the distinguished panelists and audience based on basic research findings [1, 9, 10] in the emerging field of AS:

The Key Fundamental Challenges (KFCs) to AS

1) What are the *necessary and sufficient conditions* for enabling AS?
2) Why had there rarely existed any *fully functioning AS* developed in the past 60 years?
3) Does that of KFC2 indicate a *theoretical or technical challenge*, or both?
4) How mature are our *computing platforms* for implementing AS?
5) Is *Stored-Program-Controlled* (SPC) computers, or von Neumann machines, adequate enough for designing AS? If not, what kind of computers will be needed for doing so?
6) Are our *programming languages* sufficiently expressive for implementing AS? What would happen to AS if the deterministic conditional if-then-else structures were exhausted at run-time?
7) Are our *mathematical means* ready for formally expressing AS? How would the indeterministic or unpredictable behaviors of AS be formally described in algorithms beyond classic programming languages?
8) Are typical neural-network-based *deep learning systems* an AS? Would they merely a reflexive system after training?
9) Is our *inference power* adequate for expressing real-time and indeterministic behaviors of AS?
10) How may an AS be *trusted and verified* when its state space is infinitive in de facto, such as those of self-driving vehicles and mission-critical robots?

The KFCs to AS recognized in our basic research provide a set of theoretical foundations and basic design criteria for AS development. KFCs may serve as s set of necessary and sufficient conditions for evaluating if any potential theory, methodology, solution, or implementation is suitable for AS or still yet to be. (*This section is contributed by Prof. Yingxu Wang.*)

3. AI Challenges: Knowledge Quantification, Evolution and Education in Autonomous Systems Research

In the eve of the 20th century, the famous German mathematician Hilbert issued his 23 Hilbert's problems in Mathematics. They were all unsolved in 1900 and many of them proved to be very influential for 20th-century Mathematics. Such AI challenges do exist today and, if properly defined and addressed, they can greatly boost AI and autonomous systems research [31]. Subsequently, I define three such problems: knowledge quantification, knowledge evolution/adaptation and knowledge education. Actually, it can be debated, whether they are indeed three independent problems.

Knowledge quantification. The ability of an AI/autonomous system to really operate as such depends on its knowledge of the environment and of itself (self-awareness). Unfortunately, *Knowledge* is such an elusive, yet pervasive and ubiquitous term, as it forms the basis of our society. It is found in philosophy and education texts since antiquity. Yet, the following proverb, attributed to Socrates, applies to it: 'The only thing I know is that I do not know' (εν οίδα ότι ουδέν οίδα). Even its formal definitions do not really converge. Therefore, a proper, epistemologically correct, quantifiable and practical definition of a knowledge is one of the major challenges we face today. This goes hand-in-hand with knowledge quantification.

Knowledge evolution/adaptation. Another equally pressing issue is knowledge acquisition and evolution/adaptation. In recent decade, many advances have happened in using Machine Learning for knowledge acquisition, typically in the form of Deep Neural Networks (DNNs).

Knowledge adaptation has also been addressed in a rather fragmented way, e.g., through transfer/lifelong/continual learning. Despite all this progress, major issues are still unsolved. We cannot quantify AI system (notably DNN) knowledge in a satisfactory way. As a result, we cannot quantify its evolution, when trying to learn e.g., with more/new data or new tasks. And, of course, we cannot optimize knowledge evolution. This is a major issue to be solved that will really boost system adaptation and autonomy.

Knowledge education. It defines the processes of transferring knowledge from AI system(s) and/or human(s) to other AI system(s). In this sense, its scope is much broader then the current knowledge transfer theory. Actually, I claim that the 'teacher-student' model that prevails in human education, as well as other human education theories and paradigms can be adapted to an AI education environment. Such advances can greatly boost both knowledge acquisition and knowledge evolution in AI/autonomous systems. Going the opposite way, novel Knowledge education theories can be adapted and can quantify/improve human education. (*This section is contributed by Prof. Ioannis Pitas.*)

4. Autonomous Systems: What is Missing and a Way Forward?

Despite tremendous progress and impressive results witnessed in the last decade, learning in autonomous systems is still application dependent and subject to constraints [11-12]. It is commonly assumed that real progress towards generalized artificial intelligence will only achieved when human brain inspired information processing systems are available to guide autonomous dynamic systems. Although such systems are not yet developed, desiderata are available to guide developments in the area. In the long term, a practical, human-like, information processing system should:

- Utilize a measurement module that maximizes information gain from the environment.
- Use memory-based attentional mechanisms to process information.
- Deploy a reasoning/decision making engine to identify intelligent choices in an uncertain environment.
- Rely on feedback control to interact with the environment in an efficient and cost-effective manner.

In the short term, work in this area should commence on narrow focused learning and autonomy tasks. A list of possible activities in the next two-to-five years may include but not limited to research towards the development of:

- Cognitive dynamic system for engineered autonomous systems such as robots and vehicles. Learning should focus on unsupervised learning when learning is characterized by sequential dependencies in the observations arising from the information sources (context dependency), the observation medium (perception), internal system state (structure), or the action taken (feedback loop), sequential multi-class classification, and adaptive state estimation.
- Unified processing framework which includes optimal (sub-optimal), linear (non-linear) inference for cognitive dynamic systems. Preliminary results indicate that a generalized Bayesian learning framework, with quantified risk profile, is an excellent starting point.
- Solutions that promote well-being and quality of life solutions. For example, open research problems (challenges) from the field of EEG-based brain-computer interaction (BCI) can be used to test and demonstrate the utility for such brain inspired learning framework. (*This section is contributed by Prof. Konstantinos N. Plataniotis.*)

5. Incremental Learning for Self-awareness of Autonomous Systems

Multisensor signal data fusion and perception, including processing of signals are important cognitive functionalities that can be included in artificial systems to increase their level of autonomy. However, the techniques they rely on have been developed incrementally along time with the underlying assumption that they should have been used mainly to provide a support to decision tasks driving the actions of those systems. Cognitive functionalities like self-awareness have been so far considered as not primary part of embodied knowledge of an autonomous or semiautonomous system. One of the reasons for this choice was the lack of understanding the principles that could allow an agent, even a human one, to organize successive sensorial experiences into a coherent framework of emergent knowledge, by means of integrating signal processing, machine learning and data fusion aspects. However, the developments of this last decade in many fields carried to the possibility to provide integrated solutions capable to sketch how emergent self-awareness can be obtained by capturing experiences of autonomous agents like for example vehicles and intelligent radios. In this keynote, a Bayesian approach including abnormality detection and incremental learning of generative predictive models as bricks of emergent self-awareness in intelligent agents. Discussion of the advantages of including emergent self-awareness intelligent agents will be also provided with respect to different aspects, e.g. explainability of agent's actions and capability of imitation learning. (*This section is contributed by Prof. Carlo Regazzoni.*)

6. Collaborative Autonomy is the Solution for Driverless Cars

The race to full autonomy is on, but driverless cars need to communicate and collaborate to provide for overall safety and reliability, and smart infrastructure is needed for mass adoption. This requires resilient coordination, self-healing networks, learning, and rapid collaborative decision making with humans and machines. The problem difficulty grows with environmental variation and complexity, tempo, and interaction between autonomous and human operation, while design is complicated by heterogeneity, scale, and communications rate. Interim solutions are possible in pristine or controlled environments, but widely deployed driverless cars must rely on collaboration. (*This section is contributed by Dr. Brian M. Sadler.*)

7. Learning with Limited Supervision in Autonomous Systems

The recent successes in sensing and navigation algorithms in autonomous systems have been mostly around using a huge corpus of intricately labeled data for training machine learning

models. But, in real-world cases, acquiring such large datasets will require a lot of manual annotation, which may be very time-consuming, impossible within limited resources, or even prone to errors. However, a lot of real data can be acquired at low to no annotation cost. Such data can be unlabeled or contain tag/meta-data information, termed as weak annotation. Thus, we need to develop methods that can learn recognition models from such data involving limited manual supervision. In this discussion [13-18], we look into two dimensions of learning with limited supervision - first, reducing the number of manually labeled data required to learn recognition models, and second, reducing the level of supervision from strong to weak which can be mined from the web, easily queried from an oracle, or imposed as rule-based labels derived from domain knowledge.

In the first dimension of learning with limited supervision, we will discuss how context information, often present in natural data, can be used to reduce the number of annotations required. In the second dimension - reducing the level of supervision – we will discuss how to use weak labels instead of dense strong labels, for learning dense prediction tasks. We will discuss frameworks to learn using weak labels for action detection in videos and domain adaptation of semantic segmentation models on images. All of these tasks discussed are static in nature. Continuing in the direction of learning from weak labels, we explore sequential decision-making problems, where the next input depends on the current output, e.g., in a navigation task. We look into the problem of learning robotics tasks with a small set of expert human demonstrations via decomposing the complex task into subgoals. (*This section is contributed by Prof. Amit Roy-Chowdhury.*)

8. Interaction-Centered Design for Human-Autonomy Teaming: A Strategic Perspective

The world is facing unprecedented catastrophic risks, arising from the deadly pandemics and epidemics and intersection of exponential technologies. AI and robotics as two representative technologies of the 4th Industrial Revolution continue to advance rapidly to become increasingly exploitable across domains in multiple ways. The trend raises important questions about the benefits, complications, liabilities, and risks associated with increasing autonomy in safety and mission-critical intelligent adaptive systems (IASs) [19]. IASs are human-machine symbiosis technologies that exhibit collective intelligence enabled by optimized human-machine interactions based on their joint capabilities, strengths, and limitations to achieve shared goals [19, 20].

While AI and robotics can provide solutions to a wide range of capability gaps and challenges, but the digitization of the world is not intended to replace human involvement completely. The use of AI and autonomy in IASs involves complex legal, ethical, moral, social, and cultural issues that may impede their development, evaluation, and application by their human partners as a collaborative human-autonomy symbiotic partnership [21, 22].

However, there currently exists no government policy in this regard, no coordinated approach, no organized community response, and no international research program seeking for answers to the challenge of understanding and mitigating the risks associated with operating autonomous systems [23]. Further, the lack of guidance to support the design of these IASs while keeping potential benefits, as well as limitations and potential harm, in mind. It is imperative that the appropriate and validated processes to ensure that these AI-enabled autonomous system can be used safely and effectively before they are integrated more widely into our systems, activities, operations, and society.

To support the broader applications of these advanced IAS technologies, interaction-centered design (ICD) approach has been validated and applied broadly in mission-critical systems where operators' tasks are often cognitively challenging due to the dynamic and evolving nature of operator state, task, system, and environment status. The ICD framework, its analytical techniques, design methodologies, implementation strategies, and test and evaluation processes have helped the scientific and defence communities understand the optimal means by which human operators can be teamed up with autonomy and AI to conduct missions in complex environments. The ICD approach has been recognized by NATO's Joint Capability Group Unmanned Aircraft Systems (UAS') and became a guiding principle and strategy for three Standardization Recommendations to address human-automation interaction issues. The ICD framework provided guidance on solutions to address a variety of UAS operational issues including intelligent tutoring, trust, and decision-making for weapon engagement [23-25].

This panel talk addresses broader issues when humans transmit their interactions with AI/Autonomy from "on-the-loop" to "in-the-loop" and how ICD-based approach can be applied to deliver effective human-autonomy teaming from a strategic perspective. (*This section is contributed by Dr. Ming Hou.*)

9. Trustworthiness and Cybersecurity of AS in Healthcare

The Novel Coronavirus disease (COVID-19) has abruptly and undoubtedly changed the world as we know it at the end of 2019. Given the current situation of the pandemic and predictions for the post-pandemic era, it is expected for the use of Autonomous Systems (AS) in healthcare to increase significantly in the following years. Beside pandemic effects, such an increase in dependence on AS in Healthcare can be attributed to the growing demand of heath care in rural areas and increasing needs for in-home care. In generally, AS for healthcare is not just about connected medical devices but rather an important component in the vast medical big data systems. Trustworthiness and security of AS for telemedicine and healthcare are of paramount importance as there will be an exponentially larger amount of confidential medical and personal data vulnerable to cyberattacks. Healthcare systems have recently become the most attractive attack target for cybercrimes. It is because of not only the variety, variability, and value of medical information accessible through Electronic Health Records (EHR), but also the fundamental difference between trustworthiness of AS for healthcare and other Critical Infrastructure (CI). For instance, cyber-attacks on Intensive Care Unit (ICU) respirators can immediately put human lives at harm way. It was reported by the Wall Street Journal that cyberattacks on healthcare providers and hospitals have increased to the extent that at some cases, doctors turn away patients and even some healthcare centers have completely stopped their operation due to the impossible situation to handle the post-attack

disruption. Capitalizing on the above-mentioned critical aspects of trustworthiness and cybersecurity of AS in healthcare, there is an urgent and unmet quest to examine potential cyber-attacks on healthcare AS; analyze risk liabilities and costs associated with security incidences, develop advanced AS protection and mitigation solutions. (*This section is contributed by Dr. Arash Mohammadi.*)

10. Self-awareness in Heterogeneous Multi-Robots Systems

The research field of this competition is the unsupervised anomaly detection through self-aware [26-28] autonomous systems [29], which is an active topic involving IEEE Signal Processing Society through the Autonomous Systems Initiative, and Intelligent Transportation Systems Society. The competition allows participating teams to create intelligent and autonomous unsupervised algorithms, capable of determining the normal or non-normal behavior of a ground vehicle that interplays with the environment. So, the challenge is focused to discover anomalies automatically [29] in a common dataset that is delivered to all teams who participate in the challenge.

The goal of the competition is to detect anomalies in the aerial and ground vehicles behavior based on embedded sensory data in real time and the anomalies detected by the drone camera that observes a vehicle in the surroundings. Phase one of the open competition will be designed to give teams the data sets needed to familiarize themselves with the proposed challenge. Accordingly, the provided data sets will be divided into two groups: experiments with only normal data and experiments with mixed normal and non-normal data. The data sets will be ROS based, with LiDAR, IMU and video camera-synchronized data. The students' main tasks will consist of processing the available data from the experiments containing only normal data and create/train models to differentiate between normal and non-normal data in the experiments that presented mixed information. The proposed challenge falls in the category of unsupervised learning, in which training data contains only normal instances without any anomalies, and the testing data have mixed information. Several tasks are considered, involving a ground and an aerial vehicle and different anomalies can be identified for each task.

The challenge has motivated all participants to create innovative contributions to the field of autonomous systems. Their proposed algorithms use normal known data to infer anomalies on unlabeled new data automatically. One initial step towards decision-making in autonomous systems is the understanding of the data in terms of normal or abnormal information in time series of multisensory data. The detection of anomalies is a topic that comprises several different fields, such as signal processing, intelligent systems, machine learning, and data fusion from smart sensors. It can be applied to diverse platforms and scenarios e.g., fraud detection, social media security, medical image anomaly detection, video and audio surveillance, particularly, the ICAS 2021 challenge considered autonomous ground and aerial vehicles as application cases. (*This section is contributed by Dr. Lucio Marcenaro et al.*)

11. Autonomous Surgical Robotic Systems

The overarching objective of the special track on autonomous medical robotic systems is to present new intelligent and autonomous system technologies for surgery, therapy, rehabilitation, and diagnosis that will reduce the burden on healthcare systems by making medical interventions more efficient, accurate, accessible, and reliable. Autonomy in medicine can significantly enhance medical interventions by utilizing the advantages offered by the real-time data processing and decision-making capabilities of machines. The need for such technologies, which include robotic and wearable systems, stems from the long wait times for medical interventions. This need will be exacerbated by the projected increase in the number of seniors in the coming years.

Autonomy and intelligence have attracted a great deal of interest in several industries. One of the emerging fields of autonomy is in medical robotics, when advanced surgical robotic systems or neurorehabilitation robotic systems are automatized to maximize accuracy and consistency while minimizing the cost and load on the healthcare systems. However, due to the close proximity with humans, the safety and efficacy of these systems are of paramount importance. Also, due to the physiological sources of modalities used in this technology, such as surface electromyography, the signal interpretation would require a specialized intelligence framework. In addition, due to the complexity of the medical tasks and, in general, human behavior, these technologies are challenged to operate in unstructured, uncertain, and stochastic environments. In this special track, we collect novel expert opinions through papers, and we hope to generate a comprehensive set of views discussing the current state, challenges, and future vision in the field. We believe that through the fusion of AI, control, and signal processing, autonomous medical robots can play an imperative role in the future of healthcare systems. The need for such systems is more pronounced due to the pandemic situation and where autonomous systems can play a critical role in securing the health of the patients and clients. The special track aims to also attract student papers and presenters to further enhance and promote the accelerated field of medical autonomy. (*This section is contributed by Dr. S. Farokh Atashzar et al.*)

12. Autonomous Systems: The Case of Ethics

Basically, systems that can change their behavior in response to unexpected event(s) during operation to accommodates for the new environment are called autonomous systems (AS). AS are usually managed, controlled and supervised by an individual or an establishment. AS are ubiquitous such as, just to name a few: (i) Unmanned Aerial Vehicles (UVA); (ii) Unmanned Underwater Vehicle (UUV); (iii) Intelligent Vehicles; and (vi) fake news. AS technologies are truly transformational, with potential benefits in both monetary and risk reduction. For instant, a self-driving car gathers information from its sensors network, analyzes such information to decide and executes actions to achieve a well-defined target at a near minimum cost and the shortest time possible. The rapid spread of such systems has created new ethical imperatives and challenges to the society which led to high demand on research in this field [37-42].

We will examine the development and operation of some autonomous systems and explain the consequences of their autonomous actions in relation to some ethical values including

but not limited to safety, bias, and privacy. To facilitate the deliberations in this panel, we assume that the issues of the level of automation and autonomy as well as issues of industrial autonomous systems as related to: (i) logical process execution, (ii) adaptability, (iii) self-governance and the like are resolved. Furthermore, we assume that the artificial intelligence of AS is an integral part of the system rather than unabridged. This simplification and generalization make it easier to tackle the pressing concept of AS ethical and social implications. We will expose the ethical abuses of the use of the UVA and reveal its violations to human rights and to the International Humanitarian Laws (IHL) using the interpretation of normative ethical theories characteristics. Finally, we will suggest some ethical measures for the developers and operators of the UAVs to arrive at thoughtful ethics abiding autonomous. (*This section is contributed by Prof. Saif alZahir*)

13. Conclusion

This paper has presented a summary of the plenary panel on the Future Development of Autonomous Systems in the inaugural IEEE International Conference on Autonomous Systems (ICAS 2021) held in Montreal, Canada as a virtual conference during Aug. 10-13, 2021. Ten distinguished panelists have been invited to express their visions, insights, and latest breakthroughs towards AS. Highly interesting discussions and interactions with the audience have been conducted. A Hierarchical Intelligence Model (HIM) has been introduced to explain the nature, essences, and constraints of AS in both theoretical foundations and innovative applications. A set of 10 Key Fundamental Challenges (KFCs) to AS has been explored and discussed. The expected future work to address the challenges to AS due to the lack of cognitive, intelligent, computational, and mathematical readiness have been recognized by the panel. It is noteworthy that the individual statements and opinions included in this panel summary may not necessarily be shared by all panellists.

About the Panelists

Dr. Yingxu Wang is professor of cognitive systems, brain science, software science, and intelligent mathematics. He is the founding President of International Institute of Cognitive Informatics and Cognitive Computing (I2CICC). He is FIEEE, FBCS, FI2CICC, FAAIA, and FWIF. He has held visiting professor positions at Univ. of Oxford (1995, 2018-22), Stanford Univ. (2008, 16), UC Berkeley (2008), MIT (2012), and distinguished visiting professor at Tsinghua Univ. (2019-22). He received a PhD in Computer Science from the Nottingham Trent University, UK, in 1998 and has been a full professor since 1994. He is the founder and steering committee chair of IEEE Int'l Conference Series on Cognitive Informatics and Cognitive Computing (ICCI*CC) since 2002. He is founding Editor-in-Chiefs and Associate Editors of 10+ Int'l Journals and IEEE Transactions. He is Chair of IEEE SMCS TC-BCS on Brain-inspired Cognitive Systems, and Co-Chair of IEEE CS TC-CLS on Computational Life Science. His basic research has been across contemporary science disciplines of intelligence, mathematics, knowledge, robotics, computer, information, brain, cognition, software, data, systems, cybernetics, neurology, and linguistics. He has published 600+ peer reviewed papers and 38 books/proceedings. He has presented 62 invited keynote speeches in international conferences. He has served as honorary, general, and program chairs for 39 international conferences. He has led 10+ international, European, and Canadian research projects as PI. He is recognized by Google Scholar as world top 7 in Autonomous Systems, top 1-Software Science, top 1-Cognitive Robots, top 2-Cognitive Computing, and top 1-Knowledge Science. He is recognized by Research Gate as among the world's top 2.5% scholars with a read-index 398,800+.

Prof. Ioannis Pitas (IEEE fellow, IEEE Distinguished Lecturer, EURASIP fellow) received the Diploma and PhD degree in Electrical Engineering, both from the Aristotle University of Thessaloniki (AUTH), Greece. Since 1994, he has been a Professor at the Department of Informatics of AUTH and Director of the Artificial Intelligence and Information Analysis (AIIA) lab. He served as a Visiting Professor at several Universities. His current interests are in the areas of computer vision, machine learning, autonomous systems, intelligent digital media, image/video processing, human-centred computing, affective computing, 3D imaging and biomedical imaging. He has published over 1000 papers, contributed in 47 books in his areas of interest and edited or (co-)authored another 11 books. He has also been member of the program committee of many scientific conferences and workshops. In the past he served as Associate Editor or co-Editor of 9 international journals and General or Technical Chair of 4 international conferences. He participated in 71 R&D projects, primarily funded by the European Union and is/was principal investigator in 42 such projects. Prof. Pitas leads the big European H2020 R&D project MULTIDRONE: https://multidrone.eu/. He is AUTH principal investigator in H2020 R&D projects Aerial Core and AI4Media. He is chair of the Autonomous Systems Initiative https://ieeeasi.signalprocessingsociety.org/. He is head of the EC funded AI doctoral school of Horizon2020 EU funded R&D project AI4Media (1 of the 4 in Europe). He has 32000+ citations to his work and h-index 85+ (Google Scholar).

Prof. Konstantinos (Kostas) N. Plataniotis received his B. Eng. degree in Computer Engineering from University of Patras, Greece and his M.S. and Ph.D. degrees in Electrical Engineering from Florida Institute of Technology Melbourne, Florida. Dr. Plataniotis is currently a Professor with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto in Toronto, Ontario, Canada, where he directs the Multimedia Laboratory. He holds the Bell Canada Endowed Chair in Multimedia since 2014. His research interests are primarily in the areas of multimedia and knowledge media systems image/signal processing, machine learning and adaptive learning systems, visual data analysis, and affective computing. Dr. Plataniotis is a Fellow of IEEE, Fellow of the Engineering Institute of Canada, Fellow of the Canadian Academy of Engineering/L'Academie Canadienne Du Genie, and a registered professional engineer in Ontario.

He has served as the Editor-in-Chief of the IEEE Signal Processing Letters. He was the Technical Co-Chair of the IEEE 2013 International Conference in Acoustics, Speech and Signal Processing, and he served as the inaugural IEEE Signal Processing Society Vice President for Membership (2014 -2016) and General Co-Chair for the 2017 IEEE GLOBALSIP. He served as the 2018 IEEE International Conference on Image Processing (ICIP 2018), and as General Co-Chair of the 2021 International Conference on Acoustics, Speech and Signal Processing (ICASSP21). He will be the General Chair of the 2027 IEEE

International Conference on Acoustics, Speech and Signal Processing (ICASSP2027).

Prof. Carlo S. Regazzoni obtained the M.S. and PhD degrees from University of Genova, in 1987 and 1992, respectively. Since 2005, he is full professor of Cognitive Telecommunications systems at DITEN, University of Genova, Italy. He is coordinating international Interactive and Cognitive Environment PhD courses at UNIGE since 2008. His research interests include cognitive dynamic systems, adaptive and self-aware multimodal signal processing, Bayesian machine learning, Cognitive radio. He is author of peer-reviewed papers on more than 100 international journals and 350 at international conferences. He served in IEEE Signal Processing Society in many roles, including VP conferences in 2015-2017, Italy SPS Chapter Chair, 2010-2012, IEEE AVSS SC chair 2000-2010. He was General Chair, Technical Program chair and other roles in several international IEEE conferences within his research field. He is/has been associate/guest editor of several int. journals including July 2020 special issue of the Proceedings of the IEEE on Self Awareness in Autonomous Systems.

Dr. Brian M. Sadler (IEEE Life Fellow) is a senior scientist at the US Army Research Laboratory. He is Vice-Chair of the IEEE Signal Processing Society Autonomous Systems Initiative, an IEEE Distinguished Lecturer, and has been an Associate Editor for several journals in signal processing, networking, and robotics. His research interests include collaborative autonomy and intelligent systems.

Prof. Amit Roy-Chowdhury received his PhD from the University of Maryland, College Park (UMCP) in 2002 and joined the University of California, Riverside (UCR) in 2004 where he is a Professor and Bourns Family Faculty Fellow of Electrical and Computer Engineering, Director of the Center for Robotics and Intelligent Systems, and Cooperating Faculty in the department of Computer Science and Engineering. He leads the Video Computing Group at UCR, working on foundational principles of computer vision, image processing, and statistical learning, with applications in cyber-physical, autonomous and intelligent systems. He has published over 200 papers in peer-reviewed journals and conferences. He is the first author of the book Camera Networks: The Acquisition and Analysis of Videos Over Wide Areas. He is on the editorial boards of major journals and program committees of the main conferences in his area. His students have been first authors on multiple papers that received Best Paper Awards at major international conferences, including ICASSP and ICMR. He is a Fellow of the IEEE and IAPR, received the Doctoral Dissertation Advising/Mentoring Award 2019 from UCR, and the ECE Distinguished Alumni Award from UMCP.

Dr. Ming Hou received his PhD in Human Factors from the University of Toronto, Canada in 2002. He is currently a Senior Defence Scientist at Defence Research & Development Canada (DRDC) and the Principal Authority of Human-Technology Interactions within the Department of National Defence (DND), Canada where he received the prestigious Science and Technology Excellence Award in 2020. Dr. Hou is responsible for delivering technological solutions, science-based advice, and evidence-based policy recommendations to senior decision makers within DND and the Canadian Armed Forces (CAF) and their partner organizations. He also provides advice about the investment in and application of advanced technologies and methodologies for human-machine systems requirements and for AI and Autonomy science, technology and innovation strategies to the CAF and DND. He is an Integrator of the Canadian government $1.6B IDEaS program with responsibilities for guiding national R&D activities in AI, Automation, Robotics, and Telepresence. As the Canadian National Leader and Scientific Authority during Autonomous Warrior 2018 Joint Service Exercise, Dr. Hou led Canadian research and development activities to support international collaborative projects. Dr. Hou is the Co-Chair of Human Factors Specialist Committee within NATO Joint Capability Group on Unmanned Aircraft Systems (UAS). His book: “Intelligent Adaptive Systems: An Interaction-Centered Design Perspective" provided guidance for the development of NATO STANRECs on “Human Systems Integration Guidance for UAS”, “Sense and Avoid Guidance for UAS”, and “UAS Human Factors Experimentation Guidebook”. As one of the four invited Lecturers, Dr. Hou delivered NATO Lecture Series on "UAVs: Technological Challenges, Concepts of Operations, and Regulatory Issues". Dr. Hou also serves for multiple international associations/programs as chairs or board members.

Dr. Arash Mohammadi (S’08-M’14-SM’17) is an Associate Professor with Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada. Prior to joining Concordia University and for 2 years, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. Dr. Mohammadi is a registered professional engineer in Ontario. He is Director-Membership Developments of IEEE Signal Processing Society (SPS); General Co-Chair of “2021 IEEE International Conference on Autonomous Systems (ICAS),” and Guest Editor for IEEE Signal Processing Magazine (SPM) Special Issue on “Signal Processing for Neurorehabilitation and Assistive Technologies”. He serves as Associate Editor on the editorial board of IEEE Signal Processing Letters. He was Co-Chair of “Symposium on Advanced Bio-Signal Processing and Machine Learning for Assistive and Neuro-Rehabilitation Systems” as part of 2019 IEEE GlobalSIP, and “Symposium on Advanced Bio-Signal Processing and Machine Learning for Medical Cyber-Physical Systems,” as a part of IEEE GlobalSIP’18; The Organizing Chair of 2018 IEEE Signal Processing Society Video and Image Processing (VIP) Cup, and the Lead Guest Editor for IEEE Transactions on Signal & Information Processing over Networks Special Issue on “Distributed Signal Processing for Security and Privacy in Networked Cyber-Physical Systems”. He is recipient of several distinguishing awards including the Eshrat Arjomandi Award for outstanding Ph.D. dissertation from Electrical Engineering and Computer Science Department, York University, in 2013; Concordia President’s Excellence in Teaching Award in 2018, and 2019 Gina Cody School of Engineering and Computer Science’s Research and Teaching awards in the new scholar category.

Dr. Lucio Marcenaro received the degree in electronic engineering and the Ph.D. degree in computer science and electronic engineering from Genoa University in 1999 and 2003, respectively. Since 2021, he is an Associate Professor of Telecommunications with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN), University of Genoa. He has 20 years of experience in image and video sequence analysis. He has authored about 130 technical papers on signal and video processing for computer vision.

Lucio Marcenaro is associate editor of the IEEE Transactions on Image Processing and IEEE Transactions on Circuits and Systems for Video Technology, technical program co-chair for 13th International Conference on Distributed Smart Cameras (ICDSC) and for the first IEEE International Conference on Autonomous Systems (IEEE ICAS 2021), co-organizer for the 2019 Summer School on Signal Processing (S3P), General Chair of the Symposium on Signal Processing for Understanding Crowd Dynamics. He is Senior Member of IEEE and active within the IEEE Signal Processing Italy Chapter and Director of Student Services Committee (2018-2021).

Dr. S. Farokh Atashzar is an Assistant Professor of Electrical and Computer Engineering, as well as Mechanical and Aerospace Engineering at New York University (NYU). He is also with NYU WIRELESS and NYU Center for Urban Science and Progress (CUSP). Prior to joining NYU, He was a senior scientist in the Department of Bioengineering, Imperial College London, UK, sponsored by the Natural Sciences and Engineering Research Council (NSERC) of Canada. From February 2017 to August 2018, he served as a post-doctoral scientist at Canadian Surgical Technologies and Advanced Robotics (CSTAR) center. His many awards included the highly competitive NSERC Fellowship in 2018. He was ranked among the top 5 applicants in Canada for the 2018 NSERC PDF competition in the Electrical and Computer Engineering sector. Recently he has received an NSF-RAPID-COVID award to conduct research on the topic of smart wearables for detecting health anomalies using machine intelligence for COVID-19 patients. Also, he has recently received an NSF/FDA award based on his collaboration with the medical and regulatory sectors and to generate new computational brain-muscle connectivity models to analyze the recovery process of post-stroke patients. He serves as the associate editor on several journals, including IEEE Robotics and Automation Letters (RAL), Biomedical Engineering Online (a Springer-Nature journal). He has also been actively contributing to organizing several conferences. In this regard, he has been the associate editor and publication chair for IEEE ISMR 2020-2021 and Technical Vice-Chair of IEEE International Conference on Autonomous Systems. He is also the Co-chair of IEEE Technical Committee on Telerobotics.

Professor Saif alZahir received his PhD and MS degrees in Electrical and Computer Engineering from the University of Pittsburgh, Pennsylvania and the University of Wisconsin-M in 1994 and 1984 respectively. He did his postdoc at UBC, Vancouver, BC, Canada. Dr. alZahir is involved in research in the areas of image processing, deep learning, forensic computing, data security, VLSI, Networking, Corporate Governance, and Ethics. In 2003, The Innovation Council of British Columbia, Canada, named him British Columbia Research Fellow. He authored or co-authored more than 100 journal and conference articles, contributed to two books, and 5 book chapters. He is the founder and editor-in-chief of International Journal of Corporate Governance, London, England, (2008 – present), Associate Editor, IEEE ACCESS – CTSoc. Dr. alZahir is the General Chair of the IEEE – International Conference in Image Processing 2021, Anchorage, Alaska, USA; the General Chair, of the IEEE International Symposium on Industrial Electronics, ISIE, 2022; and was the General Chair of additional three international conferences including the IEEE Nanotechnology Materials and Devices Conference (NMDC) September 2015. Finally, he served on many TPCs for international conferences.

Acknowledgement

The authors would like to thank the IEEE ICAS'21 Organization Committee, Program Committees, and the AS Initiative of IEEE Signal Processing Society. This work is supported in part by the Canadian Department of National Defence through the AutoDefence project in the IDEaS program.

References

[1] Y. Wang, A. Mohammadi, L. Marcenaro, F. Atashzar, K.N. Plataniotis, C.S. Regazzoni, I. Pitas, and A. Asif *eds*. (2021), *Proceedings of IEEE 1st International Conference on Autonomous Systems* (ICAS 2021), Montreal, Canada, Aug. 10-13, IEEE Press.

[2] Y. Wang, F. Karray, S. Kwong, K.N. Plataniotis, H. Leung, M. Hou, E. Tunstel, I.J. Rudas, L. Trajkovic, O. Kaynak, J. Kacprzyk, M.C. Zhou, M.H. Smith, P. Chen and S. Patel (2021) "On the philosophical, cognitive and mathematical foundations of symbiotic autonomous systems". *Phil. Trans. R. Soc. (A)*, 379:(2207): 1-20. https://doi.org/10.1098/rsta.2020.0362.

[3] Y. Wang, M. Hou, K.N. Plataniotis, S. Kwong, H. Leung, E. Tunstel, I.J. Rudas, and L. Trajkovic (2021), Towards a Theoretical Framework of Autonomous Systems underpinned by Intelligence and Systems Sciences, *IEEE/CAS Journal of Automatica Sinica*, 8(1), 52-63.

[4] Y. Wang, S. Yanushkevich, M. Hou, K.N. Plataniotis et al. (2020). "A Tripartite Theory of Trustworthiness for Autonomous Systems", *IEEE 2020 International Conference on Systems, Man, and Cybernetics* (SMC), Oct., IEEE Press, pp. 3375-3380.

[5] Y. Wang, S. Kwong, H. Leung, J. Lu, M.H. Smith, L. Trajkovic, E. Tunstel, K.N. Plataniotis, G. Yen, and W. Kinsner (2020), "Brain-Inspired Systems: A Transdisciplinary Exploration on Cognitive Cybernetics, Humanity, and Systems Science towards AI," *IEEE System, Man and Cybernetics Magazine*, 6(1):6–13.

[6] Y. Wang (2013), On Semantic Algebra: A Denotational Mathematics for Cognitive Linguistics, Machine Learning, and Cognitive Computing, Journal of Advanced Mathematics and Applications, 2(2), 145-161.

[7] Y. Wang (2020), Keynote: Intelligent Mathematics: A Basic Research on Foundations of Autonomous Systems, General AI, Machine Learning, and Intelligence Science, *IEEE 19th Int'l Conf. on Cognitive Informatics and Cognitive Computing* (ICCI*CC'20), Tsinghua Univ., Beijing, China, Sept., p.5.

[8] Y. Wang (2012), Inference Algebra (IA): A Denotational Mathematics for Cognitive Computing and Machine Reasoning (II), International Journal of Cognitive Informatics and Natural Intelligence, 6(1), 21-47.

[9] Y. Wang, Y. Wang, S. Patel, and D. Patel (2006), A Layered Reference Model of the Brain (LRMB), IEEE Transactions on Systems, Man, and Cybernetics (Part C), 36(2), March, 124-133.

[10] Y. Wang, S. Kwong, H. Leung, J. Lu, M.H. Smith, L. Trajkovic, E. Tunstel, K.N. Plataniotis, G. Yen, and W. Kinsner (2020). "Brain-inspired systems: A transdisciplinary exploration on cognitive cybernetics, humanity, and systems science towards AI," IEEE System, Man and Cybernetics Magazine, 6(1):6–13.

[11] M.E. Khan and H. Rue, The Bayesian Learning Rule, https://arxiv.org/abs/2107.04562, 2021.

[12] D. Farina, Arash Mohammadi, Tulay Adali, Nitish V Thakor, Konstantinos N Plataniotis, Signal Processing for Neurorehabilitation and Assistive Technologies, IEEE Signal Processing Magazine, vol. 38, no. 4, July 2021.

[13] M. Hasan, Sujoy Paul, Anastasios I Mourikis, and Amit K Roy-Chowdhury. Context-aware query selection for active learning in event recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2018.

[14] S. Paul, Jawadul H Bappy, and Amit K Roy-Chowdhury. Efficient selection of informative and diverse training samples with applications in scene classification. In IEEE Intl. Conf. on Image Processing, 2016.
[15] S. Paul, Jawadul H Bappy, and Amit K Roy-Chowdhury. Non-uniform subset selection for active learning in structured data. In IEEE Conf. on Computer Vision and Pattern Recognition, 2017.
[16] S. Paul, Sourya Roy, and Amit K Roy-Chowdhury. W-talc: Weakly-supervised temporal activity localization and classification. In European Conf. on Computer Vision, 2018.
[17] S. Paul, Yi-Hsuan Tsai, Samuel Schulter, Amit K Roy-Chowdhury, and Manmohan Chandraker. Domain adaptive semantic segmentation using weak labels. European Conf on Computer Vision, 2020.
[18] S. Paul, Jeroen Vanbaar, and Amit Roy-Chowdhury. Learning from trajectories via subgoal discovery. In Neural Information Processing Systems, 2019.
[19] M. Hou, S. Banbury, and C.Burns (2014). Intelligent adaptive systems: An interaction-centred design perspective. Boca Raton, FL: CRC Press.
[20] C. Baber (2017). Book Review: Intelligent Adaptive Systems: an interaction-centered design perspective, *Ergonomics*, vol. 60(10), 1458-1459.
[21] M. Hou (2021) Intelligent Adaptive System: A Strategic Perspective on Delivering Science Technology & Innovation on Effective Human-AI/Autonomy Integration, DRDC-RDDC-B21-0219-04748.
[22] H.M. Roff and R. Moyers (2016). Meaningful human control, artificial intelligence and autonomous weapons. Briefing paper, the Informal Meeting of Experts on Lethal Autonomous Weapon Systems, UN Convention on Certain Conventional Weapons.
[23] M. Hou, G. Ho, and D. Dunwoody (2021). IMPACTS: a trust model for human-machine teaming. Special Issue on Human-Autonomy Teaming in Military Contexts. *J. of Human-Intelligent Systems Integration*. DOI:10.1007/s42454-020-00023-x.
[24] J. Bartik et al., (2020). Autonomy strategic challenge allied IMPACT final report. *TTCP Technical Report*, TR-ASC-01-2020.
[25] M. Hou, and C.M. Fidopiastis (2016). A generic framework of intelligent adaptive learning systems: from Learning Effectiveness to Training Transfer, *Theoretical Issues of Ergonomics Science*, 18(2), 167-183.
[26] K.J. Friston, B. Sengupta, and G. Auletta (2014), "Cognitive dynamics: From attractors to active inference," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 427–445.
[27] S. Haykin and J. M. Fuster (2014), "On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 608-628.
[28] A.R. Damasio (1999), The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Harcourt Brace.
[29] C.S. Regazzoni, L. Marcenaro, D. Campo, and B. Rinner (2020), "Multisensorial generative and descriptive self-awareness models for autonomous systems," *Proceedings of the IEEE*, vol. 108, no. 7, pp. 987–1010.
[30] G. Slavic, M. Baydoun, D. Campo, L. Marcenaro, and C. Regazzoni (2021), "Multilevel anomaly detection through variational autoencoders and bayesian models for self-aware embodied agents," *IEEE Transactions on Multimedia*, p. 1.
[31] I. Pitas (2021), "AI Science and Society", in press.
[32] Y. Wang, D. Liu and G. Ruhe (2004), Formal Description of the Cognitive Process of Decision Making, *Proceedings of the 3rd IEEE International Conference on Cognitive Informatics*, IEEE CS, Press, pp. 124-130.
[33] Y. Wang (2008), On the Big-R Notation for Describing Interactive and Recursive Behaviors, *International Journal of Cognitive Informatics and Natural Intelligence*, 2(1):17-28.
[34] Y. Wang (2012), On Visual Semantic Algebra (VSA): A Denotational Mathematical Structure for Modeling and Manipulating Visual Objects and Patterns, *Software and Intelligent Sciences: New Transdisciplinary Findings*, pp.68-81.
[35] Y. Wang (2015), Concept Algebra: A Denotational Mathematics for Formal Knowledge Representation and Cognitive Robot Learning, *Journal of Advanced Mathematics and Applications*, 4(1):61-86.
[36] Y. Wang (2009), Formal Description of the Cognitive Process of Memorization, *Transactions on Computational Science*, Springer, 5(1):81-98.
[37] Edmond Awad et al, (2018), https://www.nature.com/articles/s41586-018-0637-6, *Nature*, October
[38] Tara O'Toole (2020), Remarks on "Synthetic Biology and National Security: Risks and Opportunities," Center for Strategic and International Studies, April.
[39] Manuel Müller, Timo Müller, Behrang Ashtari Talkhestani, Philipp Marks, Nasser Jazdi and Michael Weyric (2020), Industrial autonomous systems: a survey on definitions, characteristics and abilities: https://doi.org/10.1515/auto-2020-0131.
[40] T.J.M. Bench-Capon (2020), "Ethical approaches and autonomous systems" *Artificial Intelligence*, Elsevier, Jan., pp 1-15.
[41] Saif alZahir and Laura Kombo (2016), "Ethical and Legal Violation of UAVs From Computing Viewpoint", *Technical Report*, UNBC.
[42] Brijesh Dongol, Ron Bell, Ibrahim Habli, Mark Lawford, Pippa Moore, and Zeyn Saigol (2021), "Panel Discussion: Regulation and Ethics of Robotics and Autonomous Systems." *Springer Nature*, https://doi.org/10.1007/978-3-030-66494-7_14.

Advances in Autonomous Systems: A Summary of the AutoDefence Summer School at IEEE ICAS'21

Yingxu Wang, *Fellow, IEEE,* University of Calgary, AB, Canada (yingxu@ucalgary.ca)
Svetlana Yanushkevich, *SM, IEEE,* University of Calgary, AB, Canada (syanshk@ucalgary.ca)
Arash Mohammadi, *SM, IEEE,* Univ. of Concordia, Montreal, QC, Canada (arash.mohammadi@concordia.ca)
Konstantinos N. Plataniotis, *Fellow, IEEE,* University of Toronto, ON, Canada (kostas@ece.utoronto.ca)
Mark Coates, *SM, IEEE,* McGill University, Montreal, QC, Canada (mark.coates@mcgill.ca)
Baris Fidan, *SM, IEEE,* University of Waterloo, ON, Canada (fidan@uwaterloo.ca)
Marina L. Gavrilova, *SM, IEEE,* University of Calgary, AB, Canada (mgavrilo@ucalgary.ca)
Yaoping Hu, *SM, IEEE,* University of Calgary, AB, Canada (huy@ucalgary.ca)
Fakhri Karray, *Fellow, IEEE,* University of Waterloo, ON, Canada (karray@uwaterloo.ca)
Henry Leung, *Fellow, IEEE,* University of Calgary, AB, Canada (leungh@ucalgary.ca)
Ming Hou, *SM, IEEE,* Defence Research and Development Canada (ming.hou@drdc-rddc.gc.ca)

Abstract – This paper presents a panel summary on the framework of Autonomous Systems (AS) and paradigms in development. AS are advanced intelligent systems and general AI technologies triggered by the transdisciplinary development in intelligence science, system science, brain science, cognitive science, robotics, computational intelligence, and intelligent mathematics. It is recognized that, in a rigorous perspective, the only matured AS is human brains and human collective intelligence. It explains why there was rarely man-made AS in the past half century, because of the theoretical, mathematical, computational, and programming language unreadiness. Therefore, the ultimate goal of AS is to implement a brain-inspired system that may think and behave as a human counterpart in hybrid intelligent systems and general AI implementations. There is no doubt that AS will be increasingly demanded by the intelligence-based industries and societies for cognitive computers, deep machine learning systems, robotics, brain-inspired systems, mission-critical systems, self-driving vehicles, and intelligent appliances.

Keywords – *Autonomous systems, intelligence systems, general AI systems, cognitive systems, trustworthy human-machine systems*

1. Introduction

The transdisciplinary advances in intelligence, cognition, computer, cybernetic, and systems sciences have led to the emerging field of autonomous systems [1-5]. Autonomous systems address the challenges to general AI and the next generation of intelligent systems where both state spaces of their stimulus and decision-making are dynamically indeterministic at design time. These fundamental constraints to current computing and AI theories/platforms indicate why there were rarely autonomous systems being developed in the past 60 years. Therefore, the pertinacious challenges to autonomous systems are not only a technical issue, but also a theoretical demand on deep basic research towards advances theories, computing platforms, intelligent programming languages, and underpinning intelligent mathematics [2, 3, 62].

The AutoDefence Micro-Network has been established with nine Canadian universities in 2019 sponsored by the AutoDefence project of the DND IDEaS program. The research objectives of the AS consortium are to develop trustworthy technologies for autonomous human-machine systems applied in dynamic and contested defence environments. To address these challenges, the AutoDefence Micro-Network focuses on three major research themes. a) Cognitive platforms for trustworthy AS: In this theme, novel methodologies for trustworthy decision-support systems are developed; b) Cognitive models of machine learning: It focuses on autonomous learning theories for training-free AS, cognitive load modeling in contested environments, autonomous defence robots, unmanned systems, reliable target recognition, autonomous decision-support, and real-time battle-field awareness and cognition; and c) Distributed/networked autonomous systems: This theme develops innovative machine learning and signal processing solutions for attack/intrusion modeling, detection, and isolation focusing on machine autonomy, distributed cooperation, and event-based nature of man-machine teaming.

This paper presents the latest research advances of the nine member laboratories in the AutoDefence Micro-Network. A summer school of AutoDefence has been organized in the IEEE 1st International Conference of Autonomous Systems (ICAS'21) with all Co-PIs, members, associated graduated students, and industrial partners. A wide coverage is summarized in this paper that provides a theoretical framework of AS underpinned by the latest advances in intelligence, cognition, computer, and system sciences towards advanced AS.

2. Biometric Technologies for Autonomous Systems

The team of the Biometric Technologies Laboratory led by Dr. S. Yanushkevich contributes to the AutoDefence project as follows: a) We are building a framework of an intelligent Decision Support System that is a cognitive system which has

perception-action cycles as well as memory and attention mechanisms per Haykin's definition [41]; b) We are developing a taxonomical view for causal Risk-Trust-Bias inference. This approach is based on a Pearl's layered causal inference hierarchy [56, 57]; c) We further develop this concept by applying causal Bayesian networks for Risk and Trust assessment, which allows for explainable decision-making [58, 59, 60]; d) We propose to apply a variety of uncertainty measures and the corresponding models: interval, dempster-shaefer, fuzzy and subjective causal networks [61]; and e) We contribute to standardization of the Risk-Trust-Bias measures [60] based on the Admiralty Code, a NATO STANAG [6] standard. (*This section is contributed by PI, Prof. Svetlana Yanushkevich and team members.*)

3. Spatio-temporal Forecasting for Multivariate Time-series

In multi-agent autonomous systems, there are frequent occurrences where decisions are made based on predictions. For example, in planning movement through a city, the agents may take into account forecasts of the traffic intensity on various road segments in order to avoid congestion. As another example, we can consider autonomous agents cooperating to configure and manage an ad-hoc network for communications. Decisions about the configurations of deployed base stations depend upon predictions of future demand. The agents are faced with the task of forecasting future values of multiple time-series using historical values and other information (covariates). Since there is a spatial aspect to the data, the task becomes spatio-temporal forecasting. The goal is to learn the predictive relationships between different time-series considering the spatial information.

This section reviews recent advances in spatio-temporal forecasting for multivariate time-series. In settings where there are large amounts of data, deep learning approaches have started to emerge as the best performing techniques [7, 8]. In particular, graph neural networks perform well when a graph can be constructed to capture the spatial relationships [9, 10]. One particularly interesting approach involves combining more traditional model-based tracking and forecasting algorithms, such as Kalman filters and particle filters, with recurrent neural network structures [11-13]. Particle filters can perform very well if the model is well-matched to the data, but often it is difficult to specify an appropriate model. By incorporating a neural network, we can endow the algorithm with the flexibility to learn a model that is a good match to the observed data. (*This section is contributed by Co-PI, Prof. Mark Coates and team members.*)

4. Analysis and Prediction in Autonomous Biometrics Authentication Systems

Through the lens of a defense and security research, our deeply interconnected society provides a vast ground for exploration into the nature and complexity of human's nature. Abundance of online activities and communications are especially prominent from the point of view of a threat prevention and cybersecurity, where person's location, writing style or linguistic profile might assist with remote authentication or reveal emotional and psychological traits. The team of the Biometric Technologies laboratory (BTLab), led by Prof M. Gavrilova at the University of Calgary, focuses on autonomous detection of a wide range of biometric traits that might be necessary to reveal potential risks to society, to prevent cyberattacks, to improve mental health, and to protect personal privacy [14].

Biometric research is one of many domains of creative explorations, that is focused on human physical appearance, behavioral expressions, social interactions, online activities, emotions, psychology and even aesthetic preferences. The primary research is on exploring identification and risk assessment traits through advanced cognitive systems and artificial intelligence methods, including classical statistical data analytics, information fusion techniques, traditional machine learning and newly developed deep learning architectures [15, 16]. Research conducted at BTLab spans areas of biometric identity management, privacy, trustworthy decision making, multi-modal systems, artificial intelligence, big data analytics, information fusion, pattern recognition, cybersecurity, risk detection and aversion, as well as human-centered computing [17].

Biometric system is formally defined as a pattern recognition system, that can extract uniquely identifying features for subject differentiation. In a typical biometric system, features are extracted from physiological and behavioral traits. Recently, social online interactions and social data became primary avenues for cybercrime prevention and online identity disambiguation. During an authentication phase, physiological, behavioral, and social features can be extracted depending upon the availability of data; then the system is typically trained on a training subset of data and validated on the test set. The test set can be open or closed, means that new users might be enrolled into the system while it is operational. When an unknown identity is supplied to the system, identification or verification decision is made based on the matching score of the test and training sets. Confidence scores can be obtained to ensure that the error rates are low, and that the decision by the system can be trusted [18]. Earlier research conducted at the Biometric Technologies lab laid a solid foundation for intelligent processing of biometric systems, and proposed ideas of combining information fusion with biometric processing, which was published as an IGI Global book "Multimodal Biometrics and Intelligent Image Processing for Security System" [14]. This research also predicted a shift towards automated machine learning and cognitive systems within the information security domain, fueled by the emerging deep learning approaches.

Substantial number of current human-computer interaction studies include big data analytics, web browsing history studies and social network activity explorations. There are numerous applications of those fields, not only for security but also for disaster recovery, drug discovery and medical diagnostics [14]. Recently emerged domain of Social Behavioral Biometric (SBB) [16] takes advantage of user online interactions to discover unique social behavioral patterns on their own or in combination with traditional modalities such as face or gait. One of the intriguing phenomena is that social behavioral

biometrics can be extracted by observing the known behavioral biometrics (e.g., expression, interactions, gestures, voice, activities, etc.) of individuals in a specific social setting over a period of time. For instance, the idiosyncratic way of person's start to a speech can be revealed by analyzing voice data acquired from regular meetings and can act as social behavioral biometric during authentication [15]. The multi-modal system research within the context of social behavioral traits explores the major advantage of the information fusion among different biometric types: the fusion of classical physiological and behavioral data with social data and soft biometrics. The results in reliable decision-making from the biometric system, resilient to the spoof attacks, distortion or low-quality data [15]. (*This section is contributed by Co-PI, Prof. Marina Gavrilova and team members.*)

5. State-of-the-Art of Sensory Cue Integration

Approaches of artificial intelligence underlie increasing automation of human-machine systems (HMS). However, collaborative interaction between human users and machines remains essential to bring forth flexible decision-making [19]. Such interaction would ideally be constructive, should the users trust the machines to undertake intended actions and the machines provide the users adequate feedback for understanding. To maximize the efficiency of the human-machine interaction, it is thus necessary for the machines to adapt to user behaviors. Such adaptation requires optimizing machine automation for inferring the user behaviors (inputs to the machine) to achieve corresponding responses (outputs of the machine). One type of the responses is feedback, which serve as sensory cues to stimulate user senses for perception and action. Therefore, the first step of achieving the adaption is to understand how humans perceive sensory cues to fit cognition for influencing their trust.

For the first step, we focus on the integration of sensory cues, because humans naturally perceive visual, hearing, haptic (pertinent to touch), and other cues to act upon their integration. The introduction consists of two parts. One part presents the basic anatomic structure and function of the human brain for cognition [20]. This part serves a knowledge foundation related to human perception and action. Another part describes three existing models applicable for the cue integration. The models include maximum likelihood estimation [21] and proportional likelihood estimation [22] that are associated with user behaviors (e.g., task accuracy), and drift-diffusion model [23] that is related to internal processes to yield the behaviors. (*This section is contributed by Co-PI, Prof. Yaoping Hu and team members.*)

6. Deep Learning Approaches for Visual Anomaly Detection

Deep Learning has several computer vision applications in the fields such as medicine, manufacturing, security, surveillance, media, autonomous driving, and robotics etc. A truly autonomous visual system [24] should be able to handle new, unseen scenarios in which it was not trained on and is expected to considerable manage such scenarios without completely breaking down. A powerful tool towards enabling this behavior in an autonomous visual system is Anomaly Detection [25]. Anomaly Detection is a technique to identify and isolate data instances that do not conform to the defined notion of normality [26]. Several deep learning approaches have been developed over the years for image and video anomaly detection. We address the most popular methods along with few novel effective ways for architectural design. We will also discuss in detail the three major deep learning modelling paradigms for anomaly detection reconstruction-based methods, generative methods and prediction-based methods. Finally, we will discuss the emphasis on creating efficient and explainable deep learning models [27] and how they could benefit real-world applications. (*This section is contributed by Co-PI, Prof. Fakhri Karray, Prof. Baris Fidan and team members*)

7. Cognitive Radar Design using Deep Learning

Machine learning (ML) has revolutionized many scientific arenas and engineering applications. The explosive interest in ML is further accelerated when researchers discovered how to outperform humans in specific applications by deepening their networks. A new field called deep learning (DL) is born and is thriving each and every day. Nature-inspired engineering applications are not restricted to machine learning. Cognitive radar (CR) is another successful instance where a system is made smart by trying to replicate human cognition. The CR was primarily defined based on the three-element cognitive cycle: learning, reasoning, and remembering. In this light, a radar is considered cognitive when it has three main elements: 1) intelligent signal processing able to learn and gather information, 2) memory for the remembering mechanism, and 3) waveforms designed based on the gathered cognition fed back from the receiver to the transmitter. Recently, the definition of CR is further improved by considering other aspects of human cognition. In this light, different levels of radar cognition are contemplated, although not realized. The DL is regarded as a primary solution to reach these levels of cognition. This seminar presents a series of novel ideas where DL can be applied to enhancing radar cognition. First, we examine one of the most attractive challenges in cognitive radar, i.e., waveform design. Traditional approaches to this challenge are reviewed. On the DL front, we mainly focus on generative adversarial networks (GANs) and reveal how they can be applied to cognitive waveform design. Here, the generator network is the main source for waveform generation. That is, the generator is the primary interest, as most other applications for GANs. Inspiringly, this can be changed to the case where the discriminator network is of primary interest. Viewing GAN's discriminator network as a radar detector leads us to a new radar concept, with astonishing benefits, which we call adversarial radars. Here, a categorized discriminator network is trained as part of a categorized GAN, with the time-frequency scene of the received signal as the data. It then classifies the received signals according to the presents or absence of the target and other related features such as speed and location. In the end, this presentation explores this new concept in detail [28-30]. (*This section is contributed by Co-PI, Prof. Henry Leung and team members.*)

8. Autonomous and Trustworthy Connection Scheduling and Smart Contact Tracing

Recently, as a consequence of the COVID-19 pandemic, dependence on telecommunication and Smart Contact Tracing (SCT) models have significantly increased. On the one hand, preserving high Quality of Service (QoS) and maintaining low latency communication are of paramount importance. In this context, we focus on providing trustworthy access to secure communication systems with the highest achievable availability and minimum latency [31, 32]. On the other hand, there is an urgent and unmet quest to develop and design autonomous, trustworthy, and secure indoor CT solutions [33, 35]. Manual CT solutions are labor-intensive, error-prone, and time-consuming as such the focus of recent research works have been shifted towards development of autonomous and trustworthy CT models. Using advanced technologies such as Internet of Things (IoT) and high QoS communication systems autonomous CT models can recognize an infected person in close contact with others within a particular location accurately and with low latency.

This section, first, reviews Unmanned Aerial Vehicles (UAVs)-aided cellular networks [31-33] with the objective of providing high QoS for ground users in both indoor and outdoor environments. The focus will be on an ultra-dense wireless network consisting of several Femto Access Points (FAPs) and UAVs. To efficiently cope with the dynamic topology of the network and time-varying behavior of ground users, an efficient connection scheduling framework will be introduced, where ground users are autonomously trained to determine the optimal caching node. Second, we review recent advances in autonomous indoor CT solutions. While existing Global Positioning Systems (GPS) can be used to provide the required accurate localization in outdoor environments, tracking in indoors is a different and challenging task. In this regard, a newly designed Trustworthy and Blockchain-enabled Indoor Contact Tracing (TB-ICT) framework [34] will be introduced. (*This section is contributed by Co-PI, Dr. Arash Mohammadi and team members.*)

9. Visual Post-hoc Explainable AI (XAI) for Convolution Neural Networks

Over the recent past decades, Deep Neural Networks (DNN) have offered an outstanding performance in a wide variety of image recognition tasks such as image classification, instance segmentation, object detection, and etc. However, these complicated models behave like" black-boxes". Although these models generate, judging from reported in the literature results, highly accurate predictions, the rationale for their predictions are unclear for the end users. The goal of Explainable AI (XAI) is to address this shortcoming by providing meaningful, complete, understandable, and faithful explanations for the model's behavior.

For each predictive model, explanations can be produced in a wide variety of forms, depending on the consumers. XAI solutions can be divided into ad-hoc solutions that focus on training machine learning based models to provide interpretability, and post-hoc solutions that aim to provide human-friendly explanations for learned models. The focus of this tutorial presentation will be on providing post- hoc visual explanations for the Convolutional Neural Networks (CNN) trained to perform image recognition (and signal recognition) tasks.

It should be noted that XAI methodologies must satisfy certain criteria. Firstly, explanations should be easy to interpret for individual users. By looking into the explanations, a clear and sensible view of the model's perspective should be provided for the (results) consumers, whether they are familiar with any Artificial Intelligence (AI) or Machine Learning (ML) concepts. Secondly, explanations should correctly reflect the model's prediction process. XAI methods should accurately estimate how models react (perform) when certain modifications are applied to the data input. These two properties are termed as understandability and faithfulness, respectively. An XAI methodology which meets both properties can be used by engineers, researchers, and ML practitioners to ascertain the trustworthiness of the CNN predictions and outcomes.

This section introduces a systematic review of the state of the art, discuss a framework for post-hoc XAI developed at the University of Toronto and discuss its utility as trustworthiness enabler. Lastly, open research issues and future trends will be briefly discussed [36-38]. (*This section is contributed by Co-PI, Prof. K.N. Plataniotis and team members.*)

10. Cognitive Dynamic Systems - Acquiring Information from Surprise

The occurrence of uncertain, far from expected, events play a significant role in guiding the behavior of autonomous systems and biological agents. Surprise is a fundamental concept which can be used to describe various aspects of autonomous behavior, including subjectivity, expectation, interest and confidence. It drives attention, directs learning, forms memory and facilitates decision making.

This section presents a surprise minimization scheme to adaptively shift attention from processed information to new measurements in the context of system state estimation and control. It will be argued that surprise minimization is a feasible measure to express information utility and to guide novelty acquisition from data. Connections and similarities between surprise minimization learning mechanisms and frameworks are highlighted based on the principle of free energy as well as the 'negentropy' principle of information.

For demonstration purposes the talk focuses on a surprise minimization scheme which can be used to adaptively shift attention from processed information to new measurements in the context of state estimation and control. Analysis is carried out by assuming that a Gaussian noise driven linear dynamic model represents the environment. The information filter is the method of choice for estimating the state of the system. It will be shown that while the filter gains information and yields to steady state, it simultaneously, and recursively, minimizes a surprise term. The tutorial will introduce a surprise-minimization driven, multiple-model adaptive estimation, identification and decision filter. It will examine its utility as an autonomous systems model. Lastly, open research issues and future trends will be briefly discussed [39-41]. (*This section is*

also contributed by Co-PI, Prof. K.N. Plataniotis and team members.)

11. Basic Research on AS towards Autonomous Machine Learning

Autonomous Systems (AS) are a run-time deterministic intelligent system that depends not only on current stimuli or demands, but also on internal goals, historical states, cumulative acquired knowledge, and rational inferences [1]. AS enable nondeterministic behaviors to be generated at run-time as that of humans known as fully autonomous cognitive intelligence.

a) The Hierarchical Intelligence Model of AS

The emerging field of AS is triggered by interdisciplinary development in intelligence science, computer science, system science, general AI (GAI), and intelligent mathematics. A *Hierarchical Intelligence Model* (HIM) [2] is introduced to classify the levels of intelligence and their recursive properties in intelligence science based on the *abstract intelligence* (αI) theory [9] and the Layered Reference Model of the Brain (LRMB) [10] where the levels of natural and system intelligence are aggregated from those of reflexive, imperative, adaptive, autonomous, and cognitive intelligence with 16 categories of intelligent behaviors. Types of system intelligence across the HIM layers are defined by rigorous mathematical models [5].

b) Intelligent Mathematics (IM): Contemporary Mathematical Foundations of AS

It is recognized that new problems require new forms of mathematical means. The unprecedented challenges to the persistent problems in AS have been shifted from deterministic adaptive intelligence based on *stored-program control* (SPC) mechanisms to indeterministic intelligence based on machine knowledge learning and autonomous inferences [3]. These new challenges demand a set of IM [47], because almost all challenging problems in AS are identified as *hyperstructures* (H) out of the classic domain of real numbers (R). Paradigms of IM encompasses system algebra, concept algebra, semantic algebra, real-time process algebra, image-frame algebra, causal probability algebra, and inference algebra as developed in my laboratory [48, 63-66].

c) The Tripartite Theory for the Trustworthiness of AS

Trustworthiness of intelligent system in general, and AS in particular, is used to be a human perception on a black-box system embodied by its reliability and dependability. Recently, human intelligent factors underpinned by team trustworthiness have drawn much attention in AS modeling. It is revealed that trustworthiness of AS is a triparted mechanism encompassing the dimensions of structure, behavior, and system dynamic interactions known as the *Tripartite Model of AS Trustworthiness* [4]. A theoretical framework of tripartite trustworthiness has been established based on a set of rigorous mathematical models for ensuring AS-based humans and machines interactions in a hybrid and real-time environment.

The tripartite theory of AS trustworthiness presents a fundamental framework in which: a) The *to-be trust* is an assessment of the trustfulness of an entity or a structure; b) The *to-do trust* is an assessment on an action or a behavior; and c) The *system trust* is a run-time evaluation of the statistical performances and interactions between human and machine intelligence in AS. This work has been supported by case studies and experiments for proving the trustworthiness of AS and enhancing the dependability of mission-critical systems.

d) An Autonomous Semantic Learning System for Fake News Recognition

A well-known challenge to AI theories and AS technologies is autonomous fake news recognition. A piece of fake news is any syntactically, semantically, and/or sequential inconsistency in natural language expressions for statements, events, or behaviors. A novel AS for fake news recognition [49] based on machine semantic learning is implemented driven by IMs including concept algebra and semantic algebra [47, 48]. A training-free machine learning algorithm for *Differential Sentence Semantic Analyses* (DSSA) is designed and implemented for efficient detection of fake news [49]. The AS methodology and DSSA algorithm enhanced after DataCup'19 [50] have achieved a level of 70.4% accuracy that outperforms the top level of traditional data-driven neural network technologies normally at the accuracy level of 55.0%. This work has paved a way towards autonomous, training-free, and real-time trustworthy technologies for machine knowledge learning and complex semantics comprehension [51-55]. (*This section is contributed by Co-PI, Prof. Yingxu Wang and team members.*)

12. Conclusion

It has been recognized that the ultimate goal of autonomous systems is to implement brain-inspired systems that may think and behave as a human counterpart in hybrid intelligent systems. Various autonomous systems are demanded to address the challenges to classic AI technologies due to the lack of cognitive, intelligent, computational, and mathematical readiness for real-time, training-free, and mission-critical applications.

This paper has reported the panel discussions in the 2021 Summer School of the AutoDefence Micro-Network. It has summarized the latest development in AutoDefence among the nine laboratories across Canada. The coverage of this paper has presented an overview of the presentations in the Summer School of AutoDefence organized in the inaugural IEEE International Conference on Autonomous Systems (ICAS 2021).

Acknowledgment

This work is supported in part by the Canadian Department of National Defence through the AutoDefence project in the IDEaS program.

References

[1] Y. Wang, A. Mohammadi, L. Marcenaro, R.F. Atashzar, K.N. Plataniotis, C.S. Regazzoni, I. Pitas, and A. Asif *eds.* (2021), *Proceedings of IEEE 1st International Conference on Autonomous Systems* (ICAS 2021), Montreal, Canada, Aug. 10-13, IEEE Press.

[2] Y. Wang, F. Karray, S. Kwong, K.N. Plataniotis4, H. Leung, M. Hou, E. Tunstel, I.J. Rudas, L. Trajkovic, O. Kaynak, J. Kacprzyk, M.C. Zhou, M.H. Smith, P. Chen and S. Patel (2021) "On the philosophical, cognitive and mathematical foundations of symbiotic autonomous systems". *Phil. Trans. R. Soc. (A)*, 379(2207): 1-20. https://doi.org/10.1098/rsta.2020.0362.

[3] Y. Wang, M. Hou, K.N. Plataniotis, S. Kwong, H. Leung, E. Tunstel, I.J. Rudas, and L. Trajkovic (2021), "Towards a theoretical framework of autonomous systems underpinned by Intelligence and Systems Sciences," *IEEE/CAS Journal of Automatica Sinica*, 8(1), 52-63.

[4] Y. Wang, S. Yanushkevich, M. Hou, K.N. Plataniotis, M. Coates, M. Gavrilova, Y. Hu, F. Karray6, H. Leung1, A. Mohammadi, S. Kwong, E. Tunstel, L. Trajkovic, I.J. Rudas, and J. Kacprzyk (2020). "A tripartite theory of trustworthiness for autonomous systems", *2020 IEEE International Conference on Systems, Man, and Cybernetics* (SMC), Oct., IEEE Press, pp.3375-3380.

[5] Y. Wang, S. Kwong, H. Leung, J. Lu, M.H. Smith, L. Trajkovic, E. Tunstel, K.N. Plataniotis, G. Yen, and W. Kinsner (2020). "Brain-inspired systems: A transdisciplinary exploration on cognitive cybernetics, humanity, and systems science towards AI," *IEEE System, Man and Cybernetics Magazine*, 6(1):6–13.

[6] A. Poursaberi, S. Yanushkevich, M. Gavrilova, et al. (2013), "Situational awareness through biometrics," *IEEE Computer*, 46(5), pp. 102–104.

[7] R. Sen, Yu, H.-F., and Dhillon, I. S. (2019). "Think globally, act locally: A deep neural network approach to high- dimensional time series forecasting," In *Proc. Adv. Neural Info. Process. Systems*, volume 32, pp. 4837–4846.

[8] B.N. Oreshkin, Amini, A., Coyle, L., and Coates, M. J. (2021). "FC-GAGA: Fully connected gated graph architecture for spatio-temporal traffic forecasting," *In Proc. AAAI Conf. Artificial Intell*., Jan.

[9] Z. Wu, Pan, S., Long, G., Jiang, J., and Zhang, C. (2019). "Graph wavenet for deep spatial-temporal graph modeling," In *Proc. Int. Joint Conf. Artificial Intell.*, pp. 1907–1913.

[10] L. Bai, Yao, L., Li, C., Wang, X., and Wang, C.(2020), "Adaptive graph convolutional recurrent network for traffic forecasting," In *Proc. Adv. Neural Info. Process. Systems*.

[11] S. Pal, Ma, L., Zhang Y. and Coates, M. (2021). "RNN with particle flow for probabilistic spatio-temporal forecasting," in *Proc. Int. Conf. Machine Learning*, July.

[12] R. Kurle, Rangapuram, S. S., de Bézenac, E., Gu¨nnemann, S., and Gasthaus, J. (2020). "Deep rao-blackwellised particle filters for time series forecasting," *Proc. Adv. Neural Info. Process. Systems.*

[13] E. de Bézenac, Rangapuram, S. S., Benidis, K., Bohlke- Schneider, M., Kurle, R., Stella, L., Hasson, H., Gallinari, P., and Januschowski, T. (2020). "Normalizing Kalman filters for multivariate time series analysis," In *Proc. Adv. Neural Info. Process. Systems*, 2020.

[14] M. Gavrilova and Monwar M (2012). *Multimodal Biometrics and Intelligent Image Processing for Security Systems*. IGI Global.

[15] M. Gavrilova, Ahmed F, Bari H, Liu R, Liu T, Maret Y, Sieu B, Sudhakar T. (2021) Multi-modal motion capture based biometric systems for emergency response and patient rehabilitation. In *Research Anthology on Rehabilitation Practices and Therapy*, 32:653-678.

[16] M. Sultana Paul PP, Gavrilova M (2014) A concept of social behavioral biometrics: Motivation, current developments, and future trends. *Int Conf on Cyberworlds*, pp 271-278.

[17] Y. Wang, B. Widrow, A.Z. Lofti, N. Howard, S. Wood, V.C. Bhavsar, G. Budin, C. Chan, R.A. Fiorini, M. Gavrilova, D.F. Shell (2016) "Cognitive intelligence: Deep learning, thinking, and reasoning by brain-inspired systems." *IJCINI*, 10(4):1-20.

[18] Y. Wang, B. Widrow, L.A. Zadeh, N. Howard, S. Wood, V.C. Bhavsar, G. Budin, C. Chan, R. Fiorini, M. Gavrilova, D.F. Shell (2019) Cognitive Intelligence: Deep learning, thinking, and reasoning by brain-inspired system. In *Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications, Critical Explorations Series*, 84:1500-1523.

[19] R.G. Hefron et al., "Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation," *Pattern Recognition Lett.*, vol. 94, pp. 96–104, 2017.

[20] M.S. Gazzaniga, R. B. Ivry, and G. R. Mangun, Cognitive neuroscience: the biology of the mind, 4th ed. N.Y: Norton & Company, Inc, 2014.

[21] M. Rohde, L. C. J. van Dam, and M. O. Ernst, "Statistically Optimal multisensory cue integration: A practical tutorial," *Multisens Res*, vol. 29, no. 4–5, pp. 279–317, 2016.

[22] S. Tarng and Y. Hu, "Proportional likelihood estimation for integrating vibrotactile and force cues in 3D user interaction," *Proc. IEEE International Conference on Systems, Man, and Cybernetics* (SMC), Toronto, Canada, Oct. 2020, pp. 3381–3386.

[23] R. Ratcliff and G. McKoon, "The diffusion decision model: Theory and data for two-choice decision tasks," *Neural Computation*, vol. 20, no. 4, pp. 873–922, Apr. 2008.

[24] A. Johari and P. D. Swami (2020), Comparison of autonomy and study of deep learning tools for object detection in autonomous self driving vehicles", *2nd Int'l Conf. on Data, Engineering and Applications* (IDEA), pp. 1-6.

[25] G. Pang, Shen C, Cao L, Hengel AV. (2021), Deep learning for anomaly detection: A review". *ACM Computing Surveys* (CSUR). Mar., 54(2):1-38.

[26] V. Chandola, Banerjee A, Kumar V. (2009), Anomaly detection: A survey. *ACM computing surveys*. Jul., 41(3):1-58.

[27] Q. Zhang and Song-Chun Zhu (2018), Visual inter-pretability for deep learning: A survey, *frontiers of information technology electronic engineering*, p. 27-39.

[28] S.Z. Gurbuz, Griffiths HD, Charlish A, Rangaswamy M, Greco MS, Bell K. (2019). An overview of cognitive radar: Past, present, and future. *IEEE Aerospace and Electronic Systems Magazine.* Dec 1;34(12):6-18.

[29] Z. Wang, She Q, Ward TE. Generative adversarial networks in computer vision: A survey and taxonomy. *ACM Computing Surveys* (CSUR). 2021 Feb 9;54(2):1-38.

[30] H.E. Najafabadi, Henry Leung (2021), A New Concept: Adversarial Radar, and its Application to S-Band Marine Radar Systems. *DRDC contract report*, March University of Calgary.

[31] Z. Hajiakhondi-Meybodi, A. Mohammadi and J. Abouei, "Deep reinforcement learning for trustworthy and time-varying connection scheduling in a coupled UAV-based femtocaching architecture," in *IEEE Access*, vol. 9, pp. 32263-32281, 2021.

[32] Z. Hajiakhondi-Meybodi, A. Mohammadi, M. Hou, and K.N. Plataniotis (2021). "DQLEL: Deep Q-learning for energy-optimized LoS/NLoS UWB node selection," Submitted to IEEE Internet of Things (IoT) Journal.

[33] Z. HajiAkhondi-Meybodi, A. Mohammadi, J. Abouei, M. Hou, K.N. Plataniotis (2021). "Joint transmission scheme and coded content placement in cluster-centric UAV-aided cellular networks," arXiv:2101.11787, 2021.

[34] M. Salimibeniy, Z. Hajiakhondi-Meybodiz, A. Mohammadiy, and Y. Wang (2021). "A trustworthy blockchain-enabled system for indoor COVID-19 contact tracing," Submitted.

[35] Z. HajiAkhondi-Meybodi, M. Salimibeni, A. Mohammadi and K. N. Plataniotis (2021). "Bluetooth low energy and CNN-based angle of arrival localization in presence of Rayleigh fading," *IEEE*

International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 7913-7917.
[36] M. Sudhakar, Sam Sattarzadeh, Konstantinos N. Plataniotis, Jongseong Jang, Yeonjeong Jeong, Hyunwoo Kim (2021), "ADA-SISE: Adaptive semantic input sampling for efficient explanation of convolutional neural networks", *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, ICASSP21, April.
[37] S. Sattarzadeh, Mahesh Sudhakar, Konstantinos N. Plataniotis, Jongseong Jang, Yeonjeong Jeong, Hyunwoo Kim (2021), "Integrated GRAD-CAM: Sensitivity aware explanation of deep convolutional networks via integrated gradient-based scoring", *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, ICASSP21, April.
[38] S. Sattarzadeh, Mahesh Sudhakar, Anthony Lem, Shervin Mehryar, KN Plataniotis, Jongseong Jang, Hyunwoo Kim, Yeonjeong Jeong, Sangmin Lee, Kyunghoon Bae (2021), "Explaining convolutional neural networks through attribution-based input sampling and block-wise feature aggregation", *34th AAAI Conference on Artificial Intelligence*, Feb..
[39] Y. Zamiri-Jafarian and K. N. Plataniotis (2020), "Bayesian surprise in linear gaussian dynamic systems: Revisiting state estimation", *2020 IEEE International Conference on Systems, Man, and Cybernetics* (SMC), Oct., 3387-3394.
[40] S. Karush, Xiao Qi Shi, Konstantinos Plataniotis, Yuri Lawryshyn (2020), "Energy-based Surprise Minimization for Multi-Agent Value Factorization", DRLW, *34th Conference on Neural Information Processing Systems (NeurIPS)*, Dec.
[41] S. Haykin (2013), *Cognitive Dynamic Systems*, Perception-action Cycle, Radar and Radio, Cambridge Press.
[42] Y. Wang (2003), On Cognitive Computing, *Brain and Mind: A Transdisciplinary Journal of Neuroscience and Neurophilosophy*, 4(2), 151-167.
[43] Y. Wang, KN. Plataniotis, L. Marcenaro, F. Atashzar, K.N. Plataniotis, L. Pitas and A. Asif, eds (2021), Future development of autonomous systems: A report of the plenary panel at IEEE ICAS'21, *Proc. IEEE 1st Int'l Conf. on Autonomous Systems* (ICAS 2021), Montreal, Canada, Aug. 10-13, IEEE Press.
[44] Y. Wang, KN. Plataniotis, A. Mohammadi, L. Marcenaro, M. Hou, H. Leung, and M. Gavrilova (2021), Perspectives on the emerging field of autonomous systems and its theoretical framework, *Proc. IEEE 1st Int'l Conf. on Autonomous Systems (ICAS 2021), Montreal, Canada, Aug. 10-13, IEEE Press, pp. 12-18.*
[45] Y. Wang (2009), On Abstract Intelligence: Toward a unified theory of natural, artificial, machinable, and computational intelligence, *International Journal of Software Science and Computational Intelligence*, Jan., 1(1): 1-17.
[46] Y. Wang, Y. Wang, S. Patel, and D. Patel (2006), A Layered Reference Model of the Brain (LRMB), *IEEE Transactions on Systems, Man, and Cybernetics* (Part C), 36(2), March, 124-133.
[47] Y. Wang (2020), Keynote: Intelligent Mathematics (IM): A basic research on foundations of autonomous systems, general AI, machine learning, and intelligence science, *IEEE 19th Int'l Conf. on Cognitive Informatics and Cognitive Computing* (ICCI*CC'20), Tsinghua Univ., Beijing, China, Sept., p. 5.
[48] Y. Wang (2012), In search of denotational mathematics: Novel mathematical means for contemporary intelligence, brain, and knowledge sciences, *Journal of Advanced Mathematics and Applications*, 1(1), 4-25.
[49] Y. Wang and J.Y. Xu (2021), An autonomous semantic learning methodology for fake news recognition, *Proc. IEEE 1st Int'l Conf. on Autonomous Systems* (ICAS 2021), Montreal, Canada, Aug. 10-13, IEEE Press.
[50] DataCup (2019). https://www.datacup.ca/.
[51] Y. Wang (2010), Cognitive Robots: A reference model towards intelligent authentication, *IEEE Robotics and Automation*, 17(4), pp. 54-62.
[52] M. Hou, S. Banbury and C. Burns (2014), *Intelligent adaptive systems: an interaction-centered design perspective*, CRC Press, NY.
[53] Y. LeCun, Y., Y. Bengio and G.E. Hinton (2015), Deep Learning, *Nature*, 521(7553):436-444.
[54] V. Mnih et al. (2015), Human-level control through deep reinforcement learning. *Nature*, 518: 529–533.
[55] E.A. Bender (2000), *Mathematical methods in artificial intelligence*, IEEE CS Press, Los Alamitos, CA.
[56] K. Lai, S.N. Yanushkevich, V. Shmerko, M. Hou (2021), "Capturing causality and bias in human action recognition," *Pattern Recognition Letters*, issue 147, pp. 164-171.
[57] K. Lai, S.N. Yanushkevich, V. Shmerko (2021), "Intelligent stress monitoring assistant for first responders," *IEEE Access*, Issue 9, pp. 25314-25329.
[58] K. Lai, H.R.C. Oliveira, M. Hou, S.N. Yanushkevich, V. Shmerko (2020), Assessing risks of biases in cognitive decision support systems, *28th European Signal Processing Conference* (EUSIPCO), pp. 840-844.
[59] T. Truong and S.N. Yanushkevich (2020), "Detecting subject-weapon visual relationships," *IEEE Symposium Series on Computational Intelligence* (SSCI), pp. 2047-2052.
[60] K. Lai, H.R.C. Oliveira, M. Hou, S.N. Yanushkevich, V. Shmerko (2020), "Risk, trust, and bias: Causal regulators of biometric-enabled decision support," *IEEE Access*, Issue 8, pp. 148779-148792.
[61] S. C. Eastwood and S. N. Yanushkevich (2016), "Risk Assessment in Authentication Machines", In: Abielmona, R., Falcon, R., Zincir-Heywood, N., Abbass, H.A. (eds.), *Recent Advances in Computational Intelligence in Defense and Security, Series: Studies in Computational Intelligence*, vol. 621, Springer Int group, 2016.
[62] Y. Wang, I. Pitas, K.N. Plataniotis, C.S. Regazzoni, B.M. Sadler, A. Roy-Chowdhury, M. Hou, L. Marcenaro, R.F. Atashzar (2021), On Future Development of Autonomous Systems: A Report of the Plenary Panel at IEEE ICAS'21, *Proceedings of IEEE 1st International Conference on Autonomous Systems* (ICAS 2021), Montreal, Canada, Aug. 10-13, IEEE Press.
[63] Y. Wang (2008), On the Big-R Notation for Describing Interactive and Recursive Behaviors, *International Journal of Cognitive Informatics and Natural Intelligence*, 2(1):17-28.
[64] Y. Wang (2015), Concept Algebra: A Denotational Mathematics for Formal Knowledge Representation and Cognitive Robot Learning, *Journal of Advanced Mathematics and Applications*, 4(1):61-86.
[65] Y. Wang (2012), On Visual Semantic Algebra (VSA): A Denotational Mathematical Structure for Modeling and Manipulating Visual Objects and Patterns, *Software and Intelligent Sciences: New Transdisciplinary Findings*, pp.68-81.
[66] Y. Wang, D. Liu and G. Ruhe (2004), Formal Description of the Cognitive Process of Decision Making, *Proceedings of the 3rd IEEE International Conference on Cognitive Informatics*, IEEE CS, Press, pp. 124-130.

AUTHOR INDEX