IEEE/ASME MESA 2024 – 20th Int. Conference on Mechatronic, Embedded Systems and Applications
02-04 September 2024, Genova-Italy

Development of Machine Intelligence for Fully Autonomous Ground Vehicles Via Video Analysis

1st Jordon Lowe
School of Engineering and Computing
University of Central Lancashire
Preston, UK
jlowe8@uclan.ac.uk

2nd Kaya Kuru
School of Engineering and Computing
University of Central Lancashire
Preston, UK
https://orcid.org/0000-0002-4279-4166

***Abstract*—The automation of vehicles is progressing from one automation level to the next, with the goal of reaching level 5, which involves no steering wheel, pedals, brakes, or windshield. This is achieved by the vehicle taking on an increasing number of autonomous decision-making tasks under the guidance of intelligent control systems that are equipped with enhancing sensor technologies and Artificial Intelligence (AI). Major vehicle companies are competing to build the most experienced (AI-) driver on the roads. In this report, how the intelligence of Self-Driving Vehicles (SDVs) is being built by the automotive industry for the efficient deployment of handover wheels is analysed and applications of machine intelligence for SDVs are implemented through video capturing using Deep Learning (DL). The results show that i) the use of DL techniques as well as reinforcement learning (RL) – Deep RL approaches – can contribute to the intelligence of SDVs significantly and ii) SDVs, equipped with advanced mechatronics systems, can be fully autonomous with the level-5 automation as they are trained appropriately with proper datasets.**

***Index Terms*—Autonomous vehicles, self-driving vehicles, driverless vehicles, sensor fusion, autonomous driving, vehicle automation, deep learning, neural networks.**

I. INTRODUCTION

Thanks to cyber-physical systems (CPSs) and enhanced AI techniques, the ”everyday things” in our environment have become increasingly intelligent in recent years in the aspects of Automation of Everything (AoE) and Internet of Everything (IoE) [1], [2] enabling them to make decisions with an increasing degree of autonomy and little to no help from humans [3], [4]. There are many applications where the replacement of humans by machines is inevitable as intelligent machines continue to advance. In the same direction, vehicles including aerial vehicles are becoming increasingly automated by taking on more and more tasks in a diverse range of fields(e.g., logistics [5], agriculture [6], landmine detection [7]). All the big players in the automotive industry envisage a future for driverless vehicles and they have already taken notable actions within their manufacturing phases [8], [9]. By utilising the vast amount of current knowledge about sensors, actuators, telematics, and artificial intelligence (AI) acquired from the level-3 and level-4 autonomy [10], the majority of automakers hope to implement level-5 fully autonomous ground vehicles (FAGVs) on urban roadways [8]. Prototypes of autonomous vehicles (AVs) in different automation levels [8] have already been tested for millions of miles [11]. In this report, how the intelligence of Self-Driving Vehicles (SDVs) is being built by the automotive industry for the efficient deployment of handover wheels is analysed and then applications of machine intelligence for SDVs are implemented through video capturing using Deep Learning (DL), particularly, a Faster Region-based Convolutional Neural Network (RCNN).

II. VEHICLE INTELLIGENCE AND INDUSTRY

Ross [12] summarises 65 years of automotive baby steps leading to autonomous ground vehicles (AGVs), from the 1948 invention of modern cruise control to the 2030 prediction that half of all new cars will be autonomous, defying the widespread belief from a few decades ago that computers would never be able to operate a car. The major automotive players all see autonomous cars as the way of the future, and they have already made significant progress toward that goal during the manufacturing phase, with strong backing from top tech companies (e.g., Samsung [3], Intel, Nvidia, Mobileye, Microsoft) as summarised in Table I. Strictly speaking, those players are heavily investing in and experimenting with autonomous technology. It’s only a matter of time until one of them introduces the first commercially available driverless vehicle in the coming years [13], as SDVs are evolving to meet the requirements of the stakeholders (i.e. cities, governors, policymakers, legislators, organisations, manufactures, users, pedestrians, and other traffic participants) [14]. The following section gives an overview of some of the main technologies incorporated into machine intelligence. Firstly, an insight into the operational principles of current AVs is presented, with a review of the approaches used by the pioneering companies in SDVs. Next, 5G communications and their potential impact on swarm intelligence in SDVs are discussed. Further, an analysis of previously developed AI systems utilising both DL and Reinforcement Learning (RL) is given.

TABLE I: Collaborations between the automotive industry and technological companies to develop vehicle intelligence. Autonomous vehicle models on Level 3, 4 and 5 [8].

Company	Model	Level
Lyft	AV HW and SW	3, 4, 5
Google	AV SW	3, 4, 5
Audio + NVIDIA	Audio A8 [1]	3
Tesla	Tesla autopilot	3
BMW + Intel + Mobileye	VISION iNEXT [20]	3
Toyoto	Concept-i [2], Lexus [3]	3
Toyoto + Microsoft	Toyoto Edge Cases	4
Volkswagen + NVIDIA	ID Buzz [4]	4
Yandex	Yandex taxi	4
Renault	Renault Symbioz	4
Renault	Trezor	4
Rolls-Royce	103EX (customisable)	4
Volvo + Microsoft	Volvo 360c	4
Ford + Lyft	Ford Fusion	4
Chrysler + Lyft	Chrysler Pacificas	4
Alphabet + Lyft	Waymo (e.g., Koala)	4
Aptiv + Lyft	Aptiv	4
Google + Lyft	Aptiv	4
Uber + NVIDIA	Aptiv	4
Rinspeed	Rinspeed Oasis [5]	4, 5
Rinspeed	Rinspeed Σtos [6]	4, 5
Rinspeed	Rinspeed Snap [7]	5
Rinspeed	Rinspeed MicroSnap [8]	5
Mercedes-Benz	S-Class S 500 (no cockpit) [9]	5
Mercedes-Benz	F 015 Luxury in Motion [10]	5
Mercedes-Benz	Future Truck 2025 [11]	5
Lyft	Lyft	5
GM	GM	5
Uber	Uber taxi	5

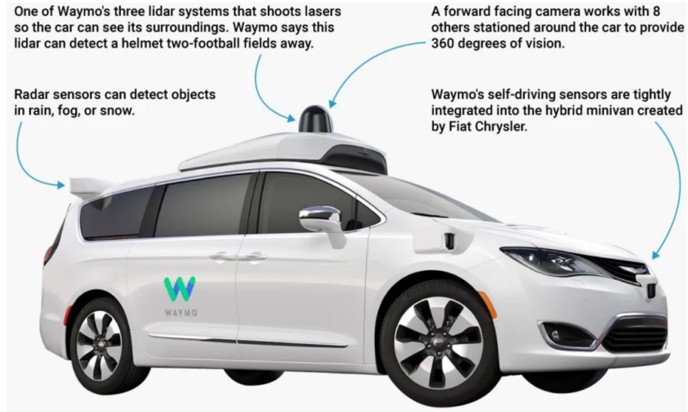

Fig. 1: Overview of Waymo's working principles [18].

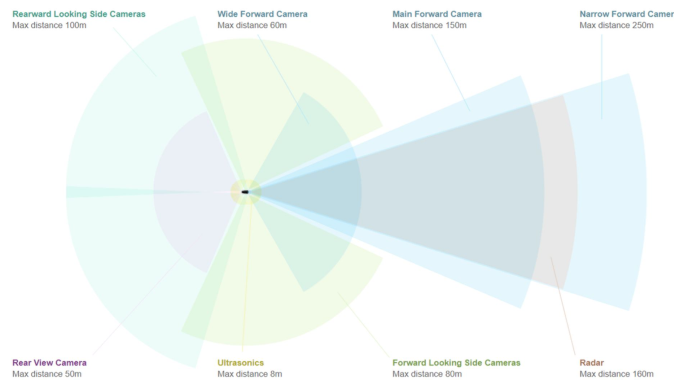

Fig. 2: Sensor fusion for Tesla's Autopilot technology [20].

A. Vehicle Intelligence With Sensor Fusion

The automotive industry spent about \$100 billion globally on research and development to foster innovation and maintain its competitiveness [15]. Around \$80 billion has been spent on self-driving cars [16]. Waymo has spent over \$1.1 billion so far to develop its own internal capability, with particularly large investments in simulation and mapping, as well as real-world driving, training, and testing [17]. The general functional architecture of these systems consists of sensors, AI for decision-making, control systems, actuators, machine-to-machine (M2M) and machine-to-infrastructure (M2I) communications.

Waymo has begun testing driverless cars on public roads massively. It, formerly Google Self Driving Car Project, has been one of the pioneering developers in AVs for the past decade. As illustrated in Fig.1, Waymo's system uses LiDAR, a laser-based distance measurement sensor, to create a detailed 360°map of the surrounding area as the vehicle drives along. Radar is also used to detect and measure the distance and velocity of nearby objects. In addition, cameras are used to recognise visual information from objects like traffic lights and signs. To enable cars to make decisions, all the gathered data is fused with machine learning methods like DL and machine vision. Tesla is another major developer of driverless cars, particularly, with level-3 automation. The main components of Tesla include traffic-aware cruise control, autosteering, auto-lane changing, side collision warnings, auto park and summoning. Tesla employs surround-view cameras to identify signs, traffic lights, and lanes in addition to radar to create an image of the objects in front of the car. GPS is used to determine the position of the car on the road, and ultrasound is also utilised in 360° to identify any obstacles. Nvidia supplies the processing power for the neural network created by Tesla that fuses all of the input data. Notably, Tesla avoids the use of LiDAR for their systems due to its high cost, complexity and obtrusive appearance [19]. Another company deeply invested in the development of driverless vehicles is Volvo. Volvo is developing both personal vehicles and vehicles such as trucks to transport large, heavy goods over short, confined areas.

B. 5G Communication and Swarm Intelligence in SDVs

The use of fifth-generation (5G) with high capacity, high speed, high data transmission rates, high reliability, high availability and low-latency abilities increases the efficacy of communication between SDVs with other vehicles and infrastructure substantially [21]. This ultra-low latency and increased transfer speeds are key aspects of 5G that enable the real-time processing and transfer of data that is required for advanced M2M and M2I communications [22], [23].

The outcome of such developments with regards to AVs would include the transfer of the vehicle's sensory data, including position, direction, intent and potential faults to surrounding vehicles and smart domains, such as Smart City (SC) or motorways to allow for proper planning and fore-warning to prevent accidents [23], [24]. Additionally, the reception of information from cloud and edge services, as well as smart domains, would be enabled to assist in route planning by the vehicle [25]. Emergency services, such as ambulances could transmit data regarding their presence to ensure other vehicles open a path, allowing quicker travel times. Other aspects include navigation, communication, efficiency, entertainment and safety [22]. These are just some of the benefits provided by enabling Vehicle-to-Everything (V2X) communications through the implementation of 5G which could be revolutionary for all modes of transport.

C. C. Deep Learning & Reinforcement Learning & Deep Reinforcement Learning

DL uses an approach like that of a human brain to process data and recognise patterns. It uses a neural network through which data is passed through each node where it is transformed by mathematical operations. This network outputs a score for each potential output, the highest score being the decision. This technique uses back-propagation to calculate the cost value (the difference between the output and the actual classification) and to adjust the weightings between the node connections to minimise this cost value [26]. DL needs substantial training which requires large datasets of known data, significant computational power and time.

An autonomous agent in RL learns by interacting with its environment using discretised state/action steps of a continuous environment (exploration) and a reward function (i.e., reward signal) to accomplish an assigned task with transition probability, T, ($< S, A, T, R >$). The agent receives a reward at each changing state naturally from the environment for desired actions (exploitation) within a goal of maximising the cumulative scalar rewards received over its lifetime for reaching the optimal policy, π^*, in other words, the highest expected sum of discounted rewards (long-term profits) during their interactions with the environment over time. Q-learning has been applied to solve various real-world problems since it was developed in [27], but it is unable to solve high-dimensional problems where the number of calculations increases drastically with the number of inputs [28] as in SDVs. Recently, Deep Neural Network (DNN) with brain-like computation has dominated the market with various application areas, in particular, computer vision, signal and natural language processing. DNN enables representation learning with artificial neural networks (NNs) using multiple layers of abstract data representations and leads to the power of generalisation in real-life applications with the efficient discovery of generalisable features in high-dimensional data.

Deep RL (DRL), with goal-directed behaviour and representation learning with the ability to learn different levels of abstraction from data, has emerged as a very effective approach by combining the strengths of two successful approaches – RL and DNN – to overcome the representation problem of RL as function approximators that leads to generalisation allowing self-driving agents to generalise knowledge to new unseen complex situations. DRL can be defined as a function approximation method in DNN to generalise past experiences to new situations with by mapping them to near-optimal decisions using generalisable optimal policies. DRL, in particular, with most commonly used Deep Q-Networks (DQN), has been found successful in addressing high dimensional problems with less prior knowledge. DRL addresses the problem of long-term optimisation within state space and time-varying environments considering Bellman‘s discounted cumulative reward through iterations. Recent revolutionary advances in AI using the learning principles of biological brains and human cognition has fuelled the development and use of DRL in numerous fields [29] such as Atari games [30], poker [31], multiplayer games [32], and board games [33], [34], [35], [36]. DRL have surpassed human-level performance in many applications. Can SDVs exceed human performance using DRL? Can self-driving with high dimensional problem space harness DRL to build systems that can learn and mimic human-driver-like problem-solving capabilities by rapidly acquiring, generalising and mapping driving knowledge to new driving tasks and situations with complex dynamics?

DRL for AVs: Autonomous driving has been suggested as an application suited to deep reinforcement learning. DRL for autonomous driving vas surveyed in [37]. Recently, DRL, using Q-learning, has been immensely employed to solve the path planning problem and maze problems of robots [38]. Vitelli and Nayebi [39] created a deep Q-network (DQN) combining convolutional and recurrent NNs (CNN & RNN) with reinforcement learning to successfully control a car in a simulated 3D environment. However, this example did prove to be less stable than other current methods. A similar approach was presented by [40] which again focused on integration of RNN & CNN, with the addition of attention models. This paper was successful in early testing. Kendall et al. [41] developed the first physical vehicle with deep reinforcement learning used to control the vehicle. The only input is a monocular image and the reward is simply the distance travelled without a safety driver intervening. The vehicle can explore continuously until the driver deems an action unsafe, this is then relayed back into the system, with all computation executed on-board. The method was able to train the vehicle to follow a lane in under 30 minutes. Kendall et al. [41] points out that although successful, a better reward system will be required, as human intervention decreases as the system improves, weakening the training input signal. With these early work, it certainly shows that DRL has some potential in AVs, as the system can dynamically adapt to new environments with extensive training or pre-built 3D mapping.

III. DEVELOPMENT OF VEHICLE INTELLIGENCE

This report shows how machine intelligence of SDVs for state and situation awareness (SSA) can be developed using visual perception. More specifically, this paper proposes a system for detecting and recognising other vehicles, their positional states, traffic signs, and road structures including lanes using DL approaches. To accomplish this, two systems have been designed. The first system uses a convolutional neural network (CNN) trained on cropped images of interested objects such as traffic signs, road signs, and vehicles. The system is used as a proof of concept for the use of DL for recognition tasks and serves as practice for the secondary system. The second system uses a Faster RCNN to detect objects in videos, draw a bounding box around them and classify their types. The following two sections detail each system.

A. Design of Machine Intelligence with CNN Detector

The CNN detector is trained on around cropped images of objects as depicted in Table II, from the German Traffic Sign Benchmark dataset [42]. To prepare the dataset, all images have been converted to grayscale and resized to 205x205 pixels. This step is used to improve the processing speed of the system. For training, the images are randomly split into 80% training and 20% validation sets. A CNN has been defined including the following layers; image input, convolutional layer, batch normalisation layer, ReLU layer and max-pooling layer. These layers are repeated twice more, with a doubled number of filters for each instance of the convolutional layer. At the end of the network, a fully connected layer, softmax layer and classification layer are used to determine the final output.

30 epochs have been determined as enough to ensure a satisfactory output accuracy, but not so much as to risk overfitting. With a trained detector, labelled test images (previously unseen by the classifier) can be passed through the system to assess its performance. Negative images (images without any object in interest) are input to evaluate the specificity of the detector. The detector will output a confidence score for each class that it is trained on, the highest scoring class being the most likely to be true, and therefore used as its determination. A threshold is set to determine the confidence value at which the detector is most accurate. A true positive (TP), false positive (FP), true negative (TN) and false negative (FN) rate are recorded through testing. A positive is any image given a value above the threshold by the classifier, and a negative image is any value below. As the images are labelled, which images are true and false is already known. These values can be used to establish the sensitivity and specificity of the system at a given threshold, using this in a ROC curve can be used to evaluate the optimal threshold value. Additionally, text-to-speech is used to audibly output the names of the objects identified by the detector.

TABLE II: Dataset used for CNN traffic sign detector.

	Traffic Sign	Training Images	Test Images
	30	2120	100
	50	2150	100
	70	1880	100
	80	1760	100
	Ahead Only	1099	100
	Keep Right	2125	100
	No Lorry Overtaking	1910	100
	Priority Road	2205	100
	Roadworks	1421	100
	Yield	2223	100
	Negatives		583

TABLE III: Dataset used for CNN traffic sign detector.

	Traffic Sign	No. of Images
	Pedestrian Crossing	1085
	Signal Ahead	925
	Stop	1820

B. Design of Machine Intelligence with Faster RCNN Detector

CNNs are well suited for image classification, however, for object detection, RCNNs (Region-CNN) can be more informative. RCNNs use selective search to generate multiple bounding boxes (called region proposals) within an image. A CNN is then used to classify the objects within the bounding boxes, these are given a score and are usually passed to a bounding box regressor to fine-tune the bounding box position [43]. Faster RCNN is an advanced RCNN that uses a region proposal network (RPN) instead of the computationally exhaustive and time-consuming selective search method [44]. The Faster RCNN used for this system uses the pre-trained ResNet-50 CNN to extract features from the input images. The extracted features are then used in the RPN to generate object proposals [45]. Finally, the proposed objects are passed to a CNN trained on the objects to determine their class. CNN is trained using the LISA dataset, which includes the region of interest (ROI) values for the bounding boxes in each image; these are used as additional input during training. The data set is split into 80% training and 20% test sets as exemplified

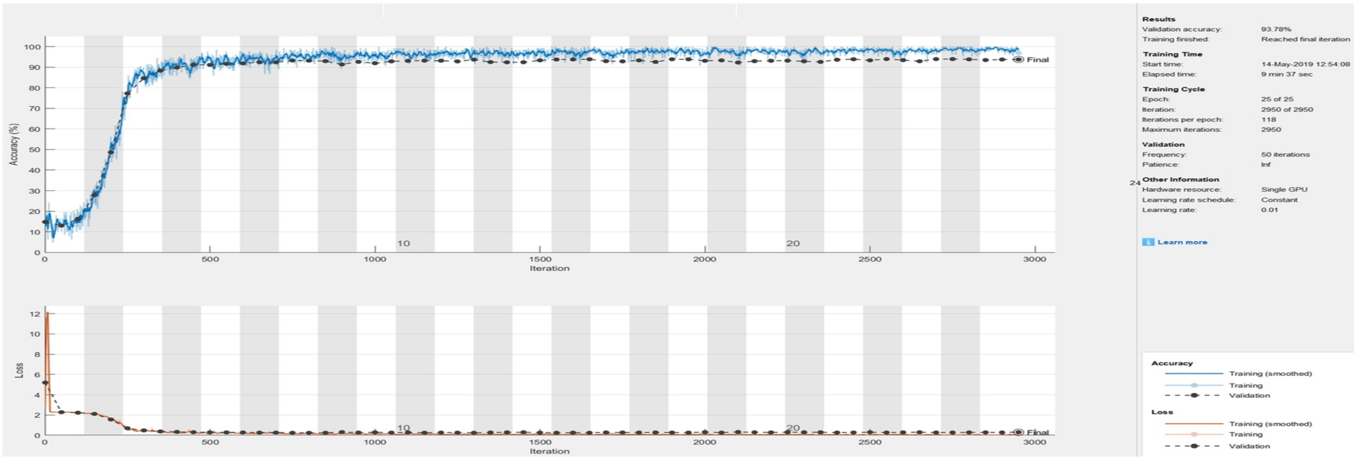

Fig. 3: Training phases of the CNN detector.

in III. The Faster RCNN is tested similarly to the CNN, again obtaining sensitivity and specificity values using various threshold values so the optimal value can be determined. The class name of the most prominent sign (highest score) is also audibly output using text-to-speech.

IV. Experimental Results

In this section, the training and testing results for both systems are presented and evaluated. Further, additional subsystems to enhance the overall system are detailed and results are given, this includes; text-to-speech, lane-detection, distance measurement and driving instructions.

A. CNN

Training the CNN using the German Traffic Sign Benchmark dataset produced a network with an accuracy of 93.78%. This accuracy was deemed suitable, as it surpassed 90%, and so was next tested using unseen images and negatives. Training phases are shown in Fig. 3. Upon testing the system, it was found that, although the sensitivity of the network was generally good, it had poor specificity. This could be attributed to several things including; poor-quality dataset images, too much similarity between trained objects or an insufficient training set. Two different negative sets were tested, the first with traffic signs other than those trained, and the second with images completely absent of signs.

Evaluation of the ROC curve generated by multiple tests found that the similarity value of 0.99 would be the ideal threshold value as shown in Fig. 4, due to it having the greatest balance between sensitivity and specificity. Overall, the system worked well for classifying images of the trained objects. The development of this system provided a good experience for working with NNs and a strong basis for developing the Faster RCNN system.

B. Faster RCNN

The LISA dataset [46] was first prepared by extracting the file path information and bounding box values for the selected

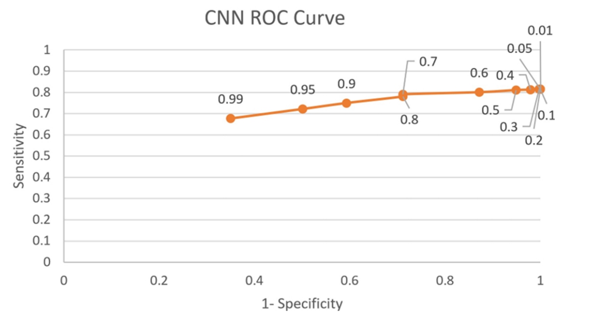

Fig. 4: The ROC curve for the CNN detector indicates that the similarity value of 0.99 would be the ideal threshold.

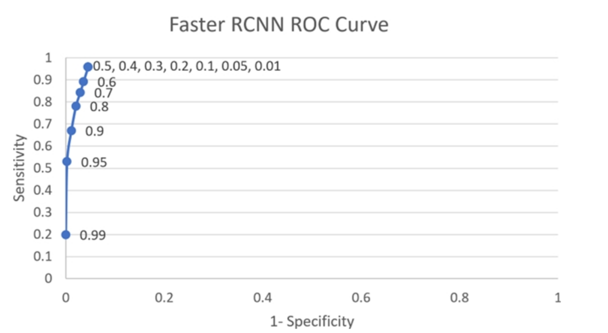

Fig. 5: ROC curve generated using multiple threshold values. Note: after 0.5, the sensitivity/specificity values did not change. This indicates a self-set threshold of 0.5 by the system. Note, data labels represent threshold value.

classes from the datasets accompanying the CSV file. This data was saved in MATLAB in a format compatible with the Faster RCNN object detector function. Initially, each class was saved to an individual set, and one detector was trained for each class. Preliminary testing showed these detectors performed well, however, using separate detectors is not satisfactory, and allows more room for errors. Consequently, the image data

Fig. 6: Faster RCNN detected and drawn bounding boxes with labels around two signal ahead signs.

Fig. 7: Faster RCNN detected and drawn bounding boxes with labels around two stop signs.

were combined into one file, with one bounding box column for each sign class. An accuracy of 100% was achieved in training this detector.

A ROC curve was generated with multiple thresholds as shown in Fig. 5. Generally, the threshold at which sensitivity and specificity are best balanced is optimal, in this case, that would be between 0.5 and 0.6 – the nearest point to the top left corner of the ROC. Preliminary tests on random images with and without the trained objects showed good results. The system was able to accurately draw bounding boxes around multiple objects of different classes, including a confidence score and class name (Figs. 6 and 7). Additionally, the system can audibly output the class name of the most prominent sign. If no sign is present, the system will produce no score and will label the image accordingly. To establish the optimal threshold value, the process used for the CNN detector was utilised. Findings indicate that the detector has a self-set threshold of 0.5, any image that scores lower will be classed as negative.

1) Testing on Video: To simulate a real-life application, the system has been tested on a video of Tesla driverless car [1]. By analysing the video frame-by-frame, the system can detect and correctly identify all signs in the video with high certainty. This proves the system will be applicable in a real vehicle. 2) Driving Information Speech Output: The system has been used for outputting driving information based on the detected objects. Firstly, a simple text-to-speech function has been incorporated into the system, which will output the class name of the object detected. This would be useful as it could inform a driver in a semi-autonomous vehicle, or passengers in a driverless vehicle. The system uses a Java file that uses natural language processing to output text-to-speech from MATLAB whilst using multiple threads. This allows the system to undertake multiple tasks simultaneously without crashing. For validation, when the vehicle in front is detected as being too close, a message will output informing the user of the car's distance and that the vehicle's speed is reducing.

[1] https://www.youtube.com/watch?v=tlThdr3O5Qo

Fig. 8: Still of driverless car video with Faster RCNN traffic sign detector implemented. This frame shows two stop signs are detected with high certainty.

3) Lane & Vehicle Detection: The trained systems has been integrated into the system that monitors the distance to the vehicle ahead, as well as the road lanes. The sensor configuration values are required to generate a birds-eye view of the road and to create a vehicle coordinate system [47] (Fig. 9). A segmentation technique is used to identify lane marker pixels from the road. This process is repeated with each frame to allow the system to hold the lane marker position. This is useful for detecting accidental lane departure and autonomous lane switching. Additionally, a pre-trained aggregated channel features (ACF) object detector is used to detect other vehicles in the frame. The vehicle coordinate system is used to calculate the location of detected vehicles. This is a critical component for allowing automatic braking and forward collision aversion. These tasks make use of multi-threaded programming to ensure they can run simultaneously with other system tasks. Two videos have been used to test the efficacy of this system, (Figs. 10 and 11).

V. Discussion and Conclusion

Full automation is anticipated to become a reality in the near future and will bring significant benefits, particularly to the transport and logistics sector [38]. The developers of the SDVs technology are claiming the safest driving on roads with the most experienced (self-) driver and promising a variety of benefits, particularly to 1) reduction of the number of accidents and traffic fatalities caused by human errors (e.g., driving

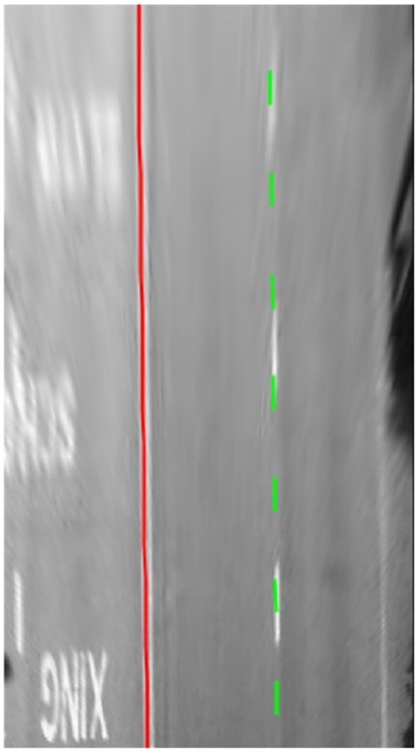

Fig. 9: Birds-eye-view of road used for lane detection.

Fig. 10: SSA in Video 1 – instant relative positioning of other vehicles with respect to the ego vehicle.

fatigue, distraction), 2) reduction of health-care costs and other financial losses caused by traffic accidents, 3) ride-hail, delivery, logistics and transportation companies for reducing their driver costs, 4) disabled and old people to make them independent (e.g., blind), 5) reduction of traffic congestion, road stress, and parking space needs, 6) people who do not like driving, and 7) comfort of passengers with optimised and standardised movement of vehicles (e.g., most convenient turning, acceleration, deceleration) [21].

This study shows guidance about how to develop the ma-

Fig. 11: SSA in Video 2 – instant relative positioning of other vehicles with respect to the ego vehicle.

chine intelligence of SDV using DL. A CNN that can classify cropped images of traffic objects showed good results when shown images of the trained objects. However, the specificity of the system is severely lacking, with a considerable number of false-positive results when shown negative images. These issues could be rectified by increasing the size and quality of the dataset. On the other hand, the Faster RCNN has shown promising results for implementation into an autonomous vehicle. The system can accurately detect and classify multiple targeted objects within an image and draw a labelled bounding box around each. Sensitivity and specificity values above 0.9 were achieved during testing. The Faster RCNN detector has been incorporated into a lane and vehicle detection system as well as the instant relative positioning of other vehicles with respect to the ego vehicle to prove that the system can successfully work as a subsystem within a driverless vehicle. Additionally, text-to-speech has been incorporated to provide driving instructions and information for passengers.

SDVs are expected to achieve certain missions within rapidly changing multi-complexity high-dimensional urban environments – partially observable, multiagent, stochastic, sequential, dynamic, and continuous environments [2]. In practice, autonomous driving consists of multiple tasks with conflicting multiple objectives. An optimal decision, with sequential decision-making, is supposed to be generated at each state/instant concerning the predictions based on the uncertain behaviours of diverse road users involving bicyclists, and pedestrians within complex urban road structures such as intersections and crosswalks. DRL, with its success in a diverse range of fields, seems a promising approach in improving the machine cognition of SDVs to mitigate the uncertainties in dynamic traffic.

References

[1] K. Kuru and H. Yetgin, "Transformation to advanced mechatronics systems within new industrial revolution: A novel framework in automation

[2]The author refers readers to [48] for the definitions of these terms.

of everything (aoe)," *IEEE Access*, vol. 7, pp. 41 395–41 415, 2019.
[2] K. Kuru, "Management of geo-distributed intelligence: Deep insight as a service (DINSaaS) on forged cloud platforms (FCP)," *Journal of Parallel and Distributed Computing*, vol. 149, pp. 103–118, Mar. 2021.
[3] ——, "Metaomnicity: Toward immersive urban metaverse cyberspaces using smart city digital twins," *IEEE Access*, vol. 11, pp. 43 844–68, 2023.
[4] K. Kuru and D. Ansell, "Tcitysmartf: A comprehensive systematic framework for transforming cities into smart cities," *IEEE Access*, vol. 8, pp. 18 615–18 644, 2020.
[5] K. Kuru, D. Ansell, W. Khan, and H. Yetgin, "Analysis and optimization of unmanned aerial vehicle swarms in logistics: An intelligent delivery platform," *IEEE Access*, vol. 7, pp. 15 804–31, 2019.
[6] K. Kuru, D. Ansell, D. Jones, B. Watkinson, J. M. Pinder, J. A. Hill, E. Muzzall, C. Tinker-Mill, K. Stevens, and A. Gardner, "Intelligent airborne monitoring of livestock using autonomous uninhabited aerial vehicles," in *The 11th European Conference on Precision Livestock Farming*, 2024.
[7] K. Kuru, D. Ansell, B. Jon Watkinson, D. Jones, A. Sujit, J. M. Pinder, and C. L. Tinker-Mill, "Intelligent automated, rapid and safe landmine and unexploded ordnance (uxo) detection using multiple sensor modalities mounted on autonomous drones," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.
[8] K. Kuru and W. Khan, "A framework for the synergistic integration of fully autonomous ground vehicles with smart city," *IEEE Access*, vol. 9, pp. 923–948, 2021.
[9] K. Kuru, "Conceptualisation of human-on-the-loop haptic teleoperation with fully autonomous self-driving vehicles in the urban environment," *IEEE Open J. Intell. Transp. Syst.*, vol. 2, pp. 448–69, 2021.
[10] ——, "Sensors and sensor fusion for decision making in autonomous driving and vehicles," 2023.
[11] R. Hussain and S. Zeadally, "Autonomous cars: Research results, issues, and future challenges," *IEEE Communications Surveys Tutorials*, vol. 21, no. 2, pp. 1275–1313, Secondquarter 2019.
[12] P. E. Ross, "Robot, you can drive my car," *IEEE Spectrum*, vol. 51, no. 6, pp. 60–90, 2014.
[13] J. Fell, "Cars of the future [transport concept cars]," *Engineering Technology*, vol. 12, no. 2, pp. 48–53, March 2017.
[14] K. Kuru, "Trustfsdv: Framework for building and maintaining trust in self-driving vehicles," *IEEE Access*, vol. 10, pp. 82 814–82 833, 2022.
[15] J. Nieuwenhuijsen, "Diffusion of Automated Vehicles A quantitative method to model the diffusion of automated vehicles with system dynamics," Master's thesis, the Delft University of Technology,, the Netherlands, 2015.
[16] L. Alter, "80 billion has been spent on self-driving cars with nothing to show for it," 2019.
[17] M. Daily, S. Medasani, R. Behringer, and M. Trivedi, "Self-driving cars," *Computer*, vol. 50, no. 12, pp. 18–23, 2017.
[18] S. Gould and D. Muoio, "Here's how waymo's brand new self-driving cars see the world," 2017.
[19] J. Hecht, "Lidar for self-driving cars," *Optics and Photonics News*, vol. 29, no. 1, p. 26, 2018.
[20] N. Heath, "Tesla's autopilot: Cheat sheet," 2018.
[21] K. Kuru, "Planning the future of smart cities with swarms of fully autonomous unmanned aerial vehicles using a novel framework," *IEEE Access*, vol. 9, pp. 6571–6595, 2021.
[22] M. Fallgren, M. Dillinger, J. Alonso-Zarate, M. Boban, T. Abbas, K. Manolakis, T. Mahmoodi, T. Svensson, A. Laya, and R. Vilalta, "Fifth-generation technologies for the connected car: Capable systems for vehicle-to-anything communications," *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 28–38, 2018.
[23] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5g: Ran, core network and caching solutions," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 4, pp. 3098–130, 2018.
[24] H. Ullah, N. Gopalakrishnan Nair, A. Moore, C. Nugent, P. Muschamp, and M. Cuevas, "5g communication: An overview of vehicle-to-everything, drones, and healthcare use-cases," *IEEE Access*, vol. 7, pp. 37 251–37 268, 2019.
[25] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5g for vehicular communications," *EEE Commun. Mag.*, vol. 56, no. 1, pp. 111–7, 2018.
[26] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436–444, May 2015.
[27] C. J. C. H. Watkins and P. Dayan, "Q-learning," in *Machine Learning*, 1992, pp. 279–292.
[28] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3826–39, 2020.
[29] K. Kuru, "Definition of multi-objective deep reinforcement learning reward functions for self-driving vehicles in the urban environment," *IEEE Trans. Veh. Technol.*, vol. 11, pp. 1–12, Mar. 2024.
[30] 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, 2015.
[31] M. Moravčík, M. Schmid, N. Burch, V. Lisý, D. Morrill, N. Bard, T. Davis, K. Waugh, M. Johanson, and M. Bowling, "DeepStack: Expert-level artificial intelligence in heads-up no-limit poker," *Science*, vol. 356, no. 6337, pp. 508–513, May 2017.
[32] M. Jaderberg, W. M. Czarnecki, I. Dunning, L. Marris, G. Lever, A. G. Castañeda, C. Beattie, N. C. Rabinowitz, A. S. Morcos, A. Ruderman, N. Sonnerat, T. Green, L. Deason, J. Z. Leibo, D. Silver, D. Hassabis, K. Kavukcuoglu, and T. Graepel, "Human-level performance in 3d multiplayer games with population-based reinforcement learning," *Science*, vol. 364, no. 6443, pp. 859–865, may 2019.
[33] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.
[34] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. P. Lillicrap, K. Simonyan, and D. Hassabis, "Mastering chess and shogi by self-play with a general reinforcement learning algorithm," *ArXiv*, vol. abs/1712.01815, 2017.
[35] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017.
[36] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play," *Science*, vol. 362, no. 6419, pp. 1140–1144, 2018.
[37] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2022.
[38] K. Kuru, J. M. Pinder, B. J. Watkinson, D. Ansell, K. Vinning, L. Moore, C. Gilbert, A. Sujit, and D. Jones, "Toward mid-air collision-free trajectory for autonomous and pilot-controlled unmanned aerial vehicles," *IEEE Access*, vol. 11, pp. 100 323–100 342, 2023.
[39] M. Vitelli and A. Nayebi, "Carma : A deep reinforcement learning approach to autonomous driving," 2016.
[40] A. E. Sallab, M. Abdou, E. Perot, and S. K. Yogamani, "Deep reinforcement learning framework for autonomous driving," in *Autonomous Vehicles and Machines*, 2017.
[41] A. Kendall, J. Hawke, D. Janz, P. Mazur, D. Reda, J.-M. Allen, V.-D. Lam, A. Bewley, and A. Shah, "Learning to drive in a day," 2018.
[42] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The german traffic sign detection benchmark," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*, 2013, pp. 1–8.
[43] S.-H. Tsang, "Review: Faster r-cnn (object detection)," 2018.
[44] VolvoTrucks, "Volvo trucks introducing vera, the future of autonomous transport," 2018.
[45] H. Gao, "Faster r-cnn explained," 2018.
[46] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1484–1497, 2012.
[47] MATHWorks, "Visual perception using monocular camera," 2019.
[48] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Saddle River, USA: Prentice Hall, 2020.

[49] W. Khan and K. Kuru, "An intelligent system for spoken term detection that uses belief combination," *IEEE Intelligent Systems*, vol. 32, no. 1, p. 70–79, Jan. 2017.

[50] K. Kuru, D. Ansell, M. Jones, B. J. Watkinson, N. Caswell, P. Leather, A. Lancaster, P. Sugden, E. Briggs, C. Davies, T. C. Oh, K. Bennett, and C. De Goede, "Intelligent autonomous treatment of bedwetting using non-invasive wearable advanced mechatronics systems and mems sensors: Intelligent autonomous bladder monitoring to treat ne," *Medical & Biological Engineering & Computing*, vol. 58, no. 5, p. 943–65, 2020.

[51] K. Kuru *et al.*, "Iotfauav: Intelligent remote monitoring of livestock in large farms using autonomous uninhabited aerial vehicles," *Computers and Electronics in Agriculture*, 2023.

[52] K. Kuru, A. Sujit, D. Ansell, J. M. Pinder, B. Jon Watkinson, D. Jones, R. Hamila, and C. Tinker-Mill, "Iintelligent, automated, rapid, and safe landmine, improvised explosive device and unexploded ordnance detection using maggy," *IEEE Access*, 2024.

[53] K. Kuru, "Platform to test and evaluate human-in-the-loop telemanipulation schemes for autonomous unmanned aerial systems," in *IEEE/ASME MESA 2024 – 20th Int. Conference on Mechatronic, Embedded Systems and Applications*, 2024.

[54] ——, "Technical report: Analysis of intervention modes in human-in-the-loop (hitl) teleoperation with autonomous unmanned aerial systems," *Central Lancashire online Knowledge*, 2024.

[55] ——, "Human-in-the-loop telemanipulation schemes for autonomous unmanned aerial systems," in *2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC)*, 2024, pp. 1–6.

[56] K. Kuru, O. Erogul, and C. Xavier, "Autonomous low power monitoring sensors," *Sensors*, vol. 21, 2021.

[57] K. Kuru, "Use of autonomous uninhabited aerial vehicles safely within mixed air traffic," in *Proceedings of Global Conference on Electronics, Communications and Networks (GCECN2024)*, 2023.

[58] ——, "Technical report: Essential development components of the urban metaverse ecosystem," *Central Lancashire online Knowledge*, 2024.

[59] ——, "Technical report: Analysis of intervention modes in human-in-the-loop (hitl) teleoperation with autonomous ground vehicle systems," *Central Lancashire online Knowledge*, 2022.

[60] ——, "Telemanipulation of autonomous drones using digital twins of aerial traffic," *IEEE Dataport*, 2024.

[61] ——, *A Novel Hybrid Clustering Approach for Unsupervised Grouping of Similar Objects*. Springer International Publishing, 2014, p. 642–653.

[62] ——, "Optimization and enhancement of h&e stained microscopical images by applying bilinear interpolation method on lab color mode," *Theoretical Biology and Medical Modelling*, vol. 11, no. 1, 2014.

[63] K. Kuru and K. Kuru, "Urban metaverse cyberthreats and countermeasures to mitigate them," in *Proceedings of IEEE Sixth International Conference on Blockchain Computing and Applications (BCCA 2024)*, 2024.

[64] ——, "Blockchain-enabled privacy-preserving machine learning authentication with immersive devices for urban metaverse cyberspaces," in *2024 20th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, 2024, pp. 1–8.

[65] K. Kuru, B. J. Watkinson, D. Ansell, D. Hughes, M. Jones, N. Caswell, P. Leather, K. Bennett, P. Sugden, C. Davies, and C. DeGoede, "Smart wearable device for nocturnal enuresis," in *2023 IEEE EMBS Special Topic Conference on Data Science and Engineering in Healthcare, Medicine and Biology*, 2023, pp. 95–96.

[66] K. Kuru, "Technical report: Big data-concepts, infrastructure, analytics, challenges and solutions," 2024.

[67] K. Kuru, D. Ansell, D. Hughes, B. J. Watkinson, F. Gaudenzi, M. Jones, D. Lunardi, N. Caswell, A. R. Montiel, P. Leather, D. Irving, K. Bennett, C. McKenzie, P. Sugden, C. Davies, and C. Degoede, "Treatment of nocturnal enuresis using miniaturised smart mechatronics with artificial intelligence," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 204–214, 2024.

[68] K. Kuru, M. Niranjan, and Y. Tunca, "Establishment of a diagnostic decision support system in genetic dysmorphology," in *2012 11th International Conference on Machine Learning and Applications*, vol. 2, 2012, pp. 164–169.

[69] J. Lowe and K. Kuru, "Design & development of a smart blind system using fuzzy logic," in *2024 20th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, 2024, pp. 1–8.

[70] K. Kuru, S. Worthington, D. Ansell, J. M. Pinder, A. Sujit, B. Jon Watkinson, K. Vinning, L. Moore, C. Gilbert, D. Jones *et al.*, "Aitl-wing-hitl: Telemanipulation of autonomous drones using digital twins of aerial traffic interfaced with wing," *Robotics and Autonomous Systems*, vol. 180, 2024.

[71] K. Kuru and K. Kuru, "Urban metaverse cyberspaces & blockchain-enabled privacy-preserving machine learning authentication with immersive devices," in *Proceedings of IEEE Sixth International Conference on Blockchain Computing and Applications (BCCA 2024)*, 2024.

[72] ——, "Urban metaverse cyberthreats and countermeasures against these threats," in *Proceedings of IEEE Sixth International Conference on Blockchain Computing and Applications (BCCA 2024)*, 2024.

[73] K. Kuru, M. Niranjan, Y. Tunca, E. Osvank, and T. Azim, "Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics," *Artificial Intelligence in Medicine*, vol. 62, no. 2, p. 105–118, Oct. 2014.

[74] K. Kuru and W. Khan, "Novel hybrid object-based non-parametric clustering approach for grouping similar objects in specific visual domains," *Appl. Soft Comput.*, vol. 62, pp. 667–701, Jan. 2018.

[75] K. Kuru, S. Clough, D. Ansell, J. McCarthy, and S. McGovern, "Wildetect: An intelligent platform to perform airborne wildlife census automatically in the marine ecosystem using an ensemble of learning techniques and computer vision," *Expert Systems with Applications*, vol. 231, p. 120574, Nov. 2023.

[76] ——, "Intelligent airborne monitoring of irregularly shaped man-made marine objects using statistical machine learning techniques," *Ecological Informatics*, vol. 78, p. 102285, Dec. 2023.

[77] K. Kuru, "Joint cognition of remote autonomous robotics agent swarms in collaborative decision-making & remote human-robot teaming," *Proceedings of The Premium Global Conclave and Expo on Robotics & Automation (AUTOROBO, EXPO2024)*, 2024.

[78] N. Caswell, K. Kuru, D. Ansell, M. J. Jones, B. J. Watkinson, P. Leather, A. Lancaster, P. Sugden, E. Briggs, C. Davies, C. Oh, K. Bennett, and C. DeGoede, "Patient engagement in medical device design: Refining the essential attributes of a wearable, pre-void, ultrasound alarm for nocturnal enuresis," *Pharmaceutical Medicine*, vol. 34, no. 1, p. 39–48, Jan. 2020.

[79] K. Kuru, D. Ansell, M. Jones, C. De Goede, and P. Leather, "Feasibility study of intelligent autonomous determination of the bladder voiding need to treat bedwetting using ultrasound and smartphone ml techniques: Intelligent autonomous treatment of bedwetting," *Medical & Biological Engineering & Computing*, vol. 57, no. 5, p. 1079–1097, Dec. 2018.

[80] K. Kuru, S. Girgin, K. Arda, and U. Bozlar, "A novel report generation approach for medical applications: The sisds methodology and its applications," *International Journal of Medical Informatics*, vol. 82, no. 5, p. 435–447, May 2013.

[81] K. Kuru and K. Kuru, "Blockchain-based decentralised privacy-preserving machine learning authentication and verification with immersive devices in the urban metaverse ecosystem," 2024.

[82] K. Kuru and S. Girgin, *A Bilinear Interpolation Based Approach for Optimizing Hematoxylin and Eosin Stained Microscopical Images*. Springer Berlin Heidelberg, 2011, p. 168–178.

[83] K. Kuru, "Use of wearable miniaturised medical devices with artificial intelligence (ai) in enhancing physical medicine," *Proceedings of Enhancing Physical Medicine. In: World Congress on Physical Medicine and Rehabilitation*, 2024.

[84] K. Kuru, A. Sujit, D. Ansell, J. M. Pinder, D. Jones, B. Watkinson, R. Hamila, and C. L. Tinker-Mill, "Non-invasive detection of landmines, unexploded ordnances and improvised explosive devices using bespoke unmanned aerial vehicles," *Proceedings of EEE International Conference on Electrical and Computer Engineering Researches (ICECER'24)*, 2024.

[85] K. Kuru, "Technical report: Towards state and situation awareness for driverless vehicles using deep neural networks," *Central Lancashire online Knowledge*, 2024.

[86] ——, "Technical report: Human-in-the-loop telemanipulation platform for automation-in-the-loop unmanned aerial systems," *Central Lancashire online Knowledge*, 2024.