

## A Multi-Modal Autonomous System for the Neural Network of Transportation Based on Recognised Driving Behaviour

Santhosh S<sup>1</sup>, M. John Basha<sup>2</sup>, I. Ambika<sup>3</sup>

<sup>1</sup>Department of CSE, Jain University (Deemed to be university, Bengaluru.

Email ID: [ambikaangel2010@gmail.com](mailto:ambikaangel2010@gmail.com)

<sup>2</sup>Department of CSE, JainUniversity(Deemed to be university, Bengaluru.

Email ID: [jjbasha@gmail.com](mailto:jjbasha@gmail.com)

<sup>3</sup>Jain University (Deemed to be university, Bengaluru.

Email ID: [ambikaangel2010@gmail.com](mailto:ambikaangel2010@gmail.com)

*Cite this paper as:* Santhosh S, M. John Basha, I. Ambika, (2025) A Multi-Modal Autonomous System for the Neural Network of Transportation Based on Recognised Driving Behaviour. *Journal of Neonatal Surgery*, 14 (11s), 378-384.

### ABSTRACT

Connected autonomous vehicles can effectively overcome the perceived limitations of human drivers through the use of communication and artificial intelligence technology. However, due to the high dynamics of the vehicular network and the various disruptions and handovers that occur, providing solid communication lines between cars is still difficult, and this might have disastrous consequences. This work proposes a technique for intelligently grouping vehicles in the heterogeneous Cognitive Internet of Vehicles based on their driving behaviours (CIoVs). The driving mode with numerous feature parameters is analysed in the proposed method to precisely capture driving traits. With the goal of facilitating trustworthy clustering of networked autonomous cars, a method based on neural network pattern recognition and the principles of evolutionary algorithms is developed. The cognitive engines can identify the different driving styles and cluster cars that share that style together. We also study the clustering mechanism's communication performance and construct the stability and life duration of clusters. Data from simulations shows that compared to state-of-the-art methods, the suggested mechanism increases reliable communication throughput by around 14.4% and average cluster lifespan by about 11.5%. lity, prolonged release, and higher antifungal activity were seen in curcumin-loaded lipid nanoparticles, suggesting that they may be an efficient antifungal delivery strategy. This method has the potential to reduce systemic toxicity and overcome drug resistance, making it a viable alternative to traditional antifungal treatments.

**Keywords:** *Cognitive Internet of Vehicles, artificial intelligence, autonomous driving, genetic algorithm, clustering mechanism*

### 1. INTRODUCTION

The purpose of the Intelligent Transportation System (ITS) is to enhance traffic management and driving efficiency, and communication and intelligence technologies contribute to this goal. Autonomous driving (AD) has attracted a lot of attention as a key component of ITS due to the positive effects it could have on traffic congestion, accident rates, and human freedom. Connected autonomous vehicles (CAVs) of the present day use vehicular communication technologies to exchange information about their surroundings and their impressions of traffic that is not in their line of sight (NLOS) [1]. These extremely competent CAVs are well suited for AD applications since they carry a wealth of computing, sensor, and networking technology. In order to make AD more secure, for instance, CAVs can work together by sharing safety alerts, perceptual data, and command information to increase security. Through concerted effort, many CAVs can be coordinated into different clusters, potentially allowing for more fuel-efficient and risk-free swarm intelligence services.



**Fig. 1 Camera image**

There is need for development in areas like mobility management and stable communication connections in vehicles, despite the fact that CIOVs have the ability to increase the intelligence and efficacy of AD services and apps. One potential issue is that the topology of the vehicle network is constantly shifting, which could lead to erratic and unreliable communication connections at first. Many applications have to be run locally or at the edge because of the high density of devices in CIOVs. A more adaptable network architecture is required in this case. Because CAVs can move from one access point to another with relative ease, communications are frequently lost and rerouted. Consistent and reliable communication is crucial to the safety of network-assisted AD, but it is challenging to ensure its availability [12-14].

Based on our findings, we suggest a new pipeline that utilizes a hybrid of lidar and camera data to detect objects that were previously invisible. Specifically, we use the language employed by Breitenstein et al. [14] to characterize corner cases at the object level—"instances that have not been observed before"—to frame our concept of anomalies. Since pedestrians and distant objects are of little concern to motorists, we will focus on those by the side of the road. Using semantic segmentation, we are able to separate the roadscape from the original image. In addition, we organize the lidar data into clusters and use a 3D object detector to look for patterns. We employ an image classifier to project 3D models into a 2D picture space, where we can then classify them. When the image classifier is unable to give an object a label, we consider it to be anomalous. Since the Waymo Open Perception Dataset [10] is the only publicly available dataset for evaluating anomalies and it includes data from more than one sensor modality, we use it for our experiments. Currently, only qualitative evaluation is possible due to a lack of ground truth data for classes other than automobiles, pedestrians, bicycles, and signage [16]. When the standards for an SAE level 3 system were finalized in December 2021, Mercedes-Benz was the first automaker in the world to achieve them. But there's a driver on board, and they need to be prepared to take the wheel at any time. However, this is not the case for autonomous systems, such as those at levels 4 and 5. Rapid progress in recent years has resulted in the widespread testing of many level 4 systems on public roads.

## 2. RELATED WORK

The majority of current anomaly detection efforts in AD contexts rely primarily on unimodal approaches. Because cameras provide more semantic context for environmental data than lidar and radar, they have seen the most development. In AD, trust-based tactics are the most effective method for detecting anomalies. They either make use of preexisting assurances in the neural network (such as class-assurance) or go through some extra difficulty to achieve the uncertainty measurement. Kendall et al. employ Monte Carlo dropout sampling to quantify the degree of uncertainty in a semantic segmentation. While others concentrate on developing learning frameworks to boost the reliability of the network. One of the goals of the training program is to mold the cutoff for appropriate action through the use of alternative methods of analysis. The latter study how to sample a distribution in feature space, which results in outliers, or how to generate outliers near the decision boundary. Each of these methods for spotting anomalies uses a cutoff value to minimize model error.

But reconstructive methods, which aim to faithfully mimic the normativity of the training data, shine when it comes to spotting outliers. If there are noticeable changes between the original and the recreation, that's a red flag. Only two groups, Ohgushi et al. and Voijr et al., have used autoencoders to re-synthesize the input picture. After that, the entropy loss from semantic segmentation and the perceptual loss between the encoder and the decoder are combined to form the anomaly map. The anomaly score, developed by Di Biase et al., incorporates chance, sensory loss, and the soft-max distance to improve upon Lis et al's re-synthesis. They employ the final segmentation output in their re-synthesis network instead of the intermediate outputs from Ohgushi et al. Instead, generative adversarial networks (GANs) and normalized flow (NFs) are used to generate odd regions or even complete driving scenarios at the training data's in/out decision boundary. The method modifies the degree of trust in the model.

## 2.1. Switching and Clustering in IoV Networks

In [26], the authors propose a section-based cluster mechanism, which organizes cars into sections of road and chooses the CH based on proximity to the cluster center rather than the vehicles' relative stability. In order to increase the network's scalability, the authors of proposed a dynamic cluster adjustment technique in which CH would dynamically modify the size of clusters depending on the availability of spectrum and the discovery of malicious cluster member (CM) nodes. To provide a seamless handoff across many heterogeneous vehicle networks, the authors of [29] describe an online reinforcement learning system to monitor data flow patterns along the space-time axis. The issue of communication switching due to the IoV's dynamic underlying network design was studied from a clustering perspective by the research team.

## 2.2. Vehicle Clustering and Driving Behaviour Modeling

Wang et al. used trajectories from linked vehicles to categorize the conditions under which drivers contact one another. Using data from many GPS traces, an extensive unsupervised learning framework can categorize the circumstances of traffic encounters. The feature representation and clustering layers make up this architecture. The researchers in used k-nearest neighbor to construct IoV by network-clustering vehicle trajectories generated via machine learning. In order to examine urban mobility patterns and quantify trajectory similarity, the authors of used density-based spatial clustering of applications and a noise technique. In , the authors presented a social-relationship-based clustering approach to ITS that takes into account both historical and real-time traffic patterns. Also demonstrated was a system for selecting routes for vehicles in traffic based on the results of an evolutionary game. Using behavioral theories and field theory, the technique developed by Tan et al. predicts driver behavior consistently across contexts. To classify drivers into one of five categories, Shahverdy et al. suggested mining vehicle data with a convolutional neural network. There was no optimization or analysis of network throughput in a linked automated driving scenario, although these publications did look at vehicle behavior in a single traffic problem and cluster vehicles in physical space.

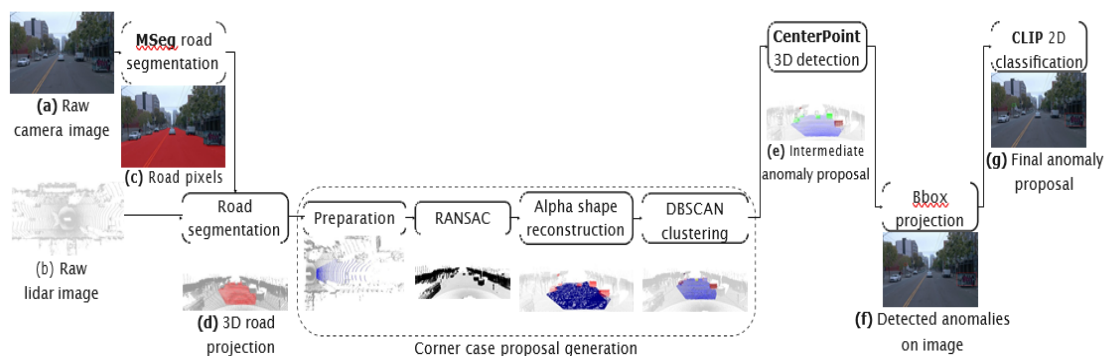
## 3. ARCHITECTURE AND CLUSTERING MECHANISM

We propose a new pipeline that integrates lidar and camera data for the discovery of unknown objects. To isolate the crucial driving area, we apply semantic segmentation to the input image. We utilize this data to generate a road mask, which we then use to overlay the 3D lidar landscape. We use a technique called clustering, which may be thought of as the strongest evidence that an object belongs to a particular class. We then compare these results to those obtained by a 3D object detector. The higher-information 2D picture space is mapped onto the grouped points of unrecognized objects. An image classifier is employed in this scenario. Starting from the premise that an object is only regarded anomalous if the associated image classifier fails to correctly identify it, we present in detail our anomaly detection workflow.

We also provide a comprehensive breakdown of the suggested architecture and the AD-driven intelligent clustering method. We begin with a high-level description of the proposed architecture for the heterogeneous CIOV, including an explanation of the roles played by each of the suggested components. We next proceed to describe AD patterns and construct a pattern recognition training model. Finally, we describe how we choose and classify CHs using a pattern-recognition-based method.

### 3.1 Road Segmentation

Specifically, we used a semantic segmentation model in the picture domain to help us zero in on roadside items. We took use of the enhanced information density offered by the camera data by using the shape and texture information.



contrasted with the sparse lidar point cloud in terms of characteristics . Waymo's Open Dataset is not used to train or test semantic segmentation models since it does not contain labelled data at the pixel level. As a result, we relied on a model that had already been trained using the MSeg dataset . MSeg is a compilation of different semantic segmentation groundtruth datasets, including COCO , ADE20K [22], Mapillary , IDD [21], BDD100K , Cityscapes [23], and SUN RGBD . Modeling with HRNet-W48 design leads to improved generalisation to new data. As a benchmark "for assessing the resilience of

models trained on different datasets," WildDash places top. An mIoU of 63.5% is achieved on BDD100K, and 76.3% is achieved on Cityscapes. We take each image and extract the road pixels, then project them into the 3D lidar space. That way, we may get a road mask and use it to filter out the irrelevant details and focus on the important ones.

### 3.2 3D Lidar Detection

CenterPoint++ was initially used to assign labels to known objects in the 3D lidar environment. Using a mAPH of 72.8 and an inference speed of 57.1ms, CenterPoint++ placed second in the Waymo Real-time 3D detection competition [10][19][20] CenterPoint++ was selected because the model relies exclusively on lidar data and its implementation is freely available, in addition to its superior performance. A two-staged model architecture based on VoxelNet was employed to achieve this precision. The current and preceding two-point cloud frames are aggregated in a multi-sweep and used as an input to the model. Finally, when at least half of the points in a cluster of objects fit within CenterPoint++'s bounding box detection, we marked the cluster as known. In 3D lidar scans, the remaining unseen objects are flagged as outliers.

### 3.3 2D camera detection

In order to improve performance beyond just delivering the labels, a task-specific cue, such as "A photo of a label," must be developed. Without the need for pre-training on a specific dataset, CLIP's zero-shot capabilities beat fully supervised models across a wide range of benchmarks and tasks. On ImageNet zeroshot, for instance, CLIP is as accurate as a pre-trained ResNet50. Since CLIP does not need any pre-training to achieve convincing zeroshot accuracy and the labels are open to interpretation, we are not limited to a tiny range of predefined object classes. We can more effectively achieve our goal of anomaly detection by specifying classes of elements that generally present in a driving environment as our nonanomaly classes than if we were limited to utilizing the labels of datasets like Waymo, which only use automobiles, pedestrians, bicycles, and signs.

Instead of just giving the labels, a task-specific prompt like "A photo of a label" can be engineered to boost performance. In contrast to fully supervised models, CLIP's zero-shot capacity is comparable across a variety of benchmarks and tasks without the need for pre-training on a specific dataset. On ImageNet zeroshot, for instance, CLIP achieves results on par with those of a pre-trained ResNet50. Since the labels for CLIP are completely up for grabs and no pre-training is required to achieve convincing zeroshot accuracy, we are not limited to any particular set of item categories. Instead of being limited to the labels of datasets like Waymo, which only employ automobiles, pedestrians, cyclists, and road signs, we are able to accomplish our goal of anomaly detection by designating classes of items that frequently exist in a driving situation as our nonanomaly classes.

Our choice of "A snapshot of a label on a street" as our cue was based on its greater specificity to the task at hand. The findings also show that CLIP is considerably more reliable than other supervised ImageNet models of comparable accuracy. However, when training on only one dataset, accuracy drops significantly on other datasets. However, because they do not rely on these correlations, zero-shot models like CLIP are more robust to shifts in distribution. Despite the challenging environment and the large variation in item size, CLIP is successful in our situation because of its generalization abilities and robustness.

### 3.4 CIOVs Architecture for AD

It has been shown that the prediction probability from a softmax distribution does not directly equate to the confidence of the model, making it difficult to utilize a threshold of the classifier's prediction probability to identify an anomaly. Only calibrated models can have their confidence deduced directly from the output prediction score. However, we discovered that the prediction probability of outliers, or anomalies, is generally lower than that of examples falling within the distribution, or non-anomalies. In addition, CLIP has been shown to be accurately calibrated on both normal and atypical distribution data. In this way, the threshold technique can serve as a reliable reference point for identifying outliers.

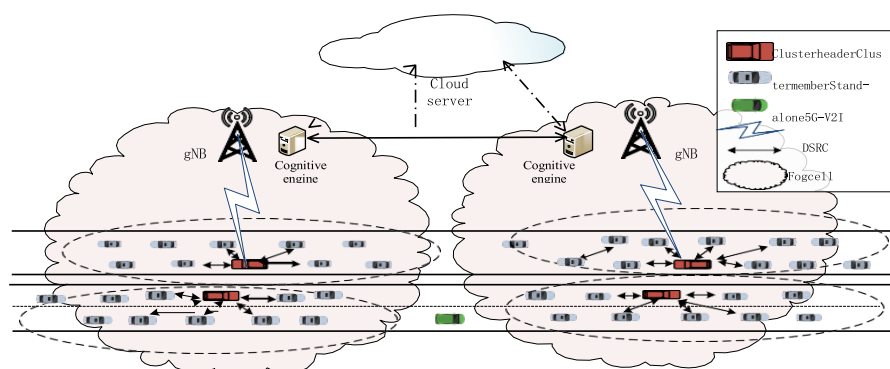


Figure 4. Heterogeneous CIOVs architecture.

(1) CAVs: Clustered vehicles (CHs), swarm vehicles (CMs), and freestanding vehicles (SAs) are the three broad categories of CAVs. To be more specific, DSRC is an efficient V2V communication system that facilitates cluster-wide communication for the purpose of delivering time-critical safety services. The SA CAVs have a direct line of communication with the control centre.

(2) gNB: The gNB is a network node that coordinates and controls all communications with both CHs and SA vehicles. To expand coverage and access to the internet beyond the confines of the cluster, the CH can, for instance, establish a connection with the gNB through 5G-V2I technology, which can then use the preexisting communication infrastructure to connect to the CS.

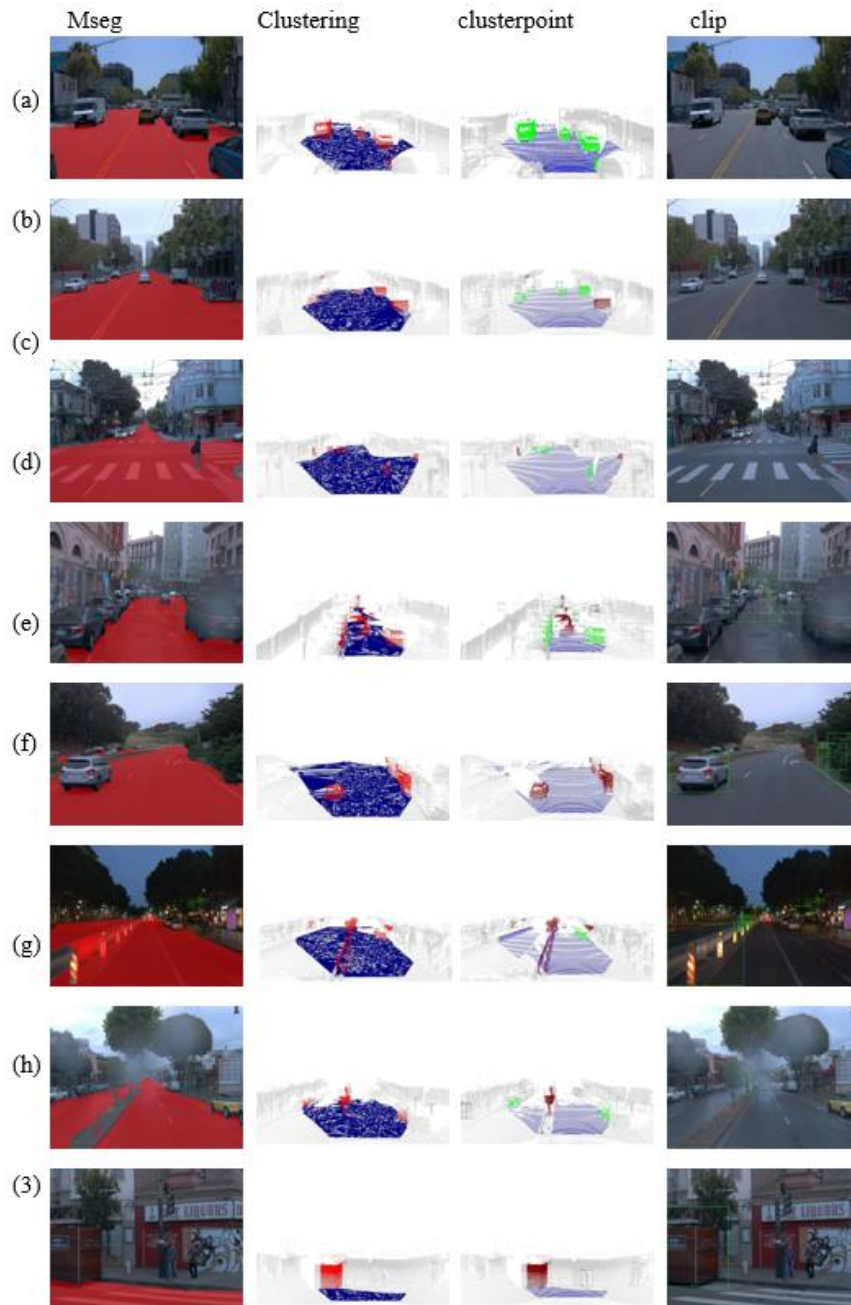


Fig 6. Res mages

CE: With the use of context and position data, the CE can recognise AD modes and move between them to maintain cluster stability and maximise driving economy. When a new mode becomes available, for instance, the driver can switch between

them as necessary. Fourthly, CS: the cloud server may provide worldwide traffic information and services by retrieving delay-insensitive data from the gNB. We use DSRC and 5G-V2I as our primary communication technology to carry out our various communication services.

The total number of CAVs (V) in a fog cell is denoted by the formula  $V = V_1, V_2, V_3, \dots, V_N$ , where N is the number of vehicles. CAVs produce copious amounts of FCD and networking data in the fog cell. Promotion of ITS and network-assisted AD can benefit from processing this data to gain useful insights. Based on the data we uncovered about their preferred driving styles, we may classify these UAVs into several groups, denoted by the letters  $C = C_1, C_2, \dots, C_L$ . The vehicles in the fog cell can switch between three distinct AD modes, and this spectrum of approaches can be represented as  $V\_ID = M_1, M_2, M_3$ , where M1 stands for a cautious driving style, M2 for an intermediate one, and M3 for an aggressive one. Thus, the vehicle's category attribute can be written as  $V_i\_v\_V\_ID$ . In the following section, we will discuss the basics of AD mode modeling.

FCD is gathered by the gNB and sent on to the CE, which then analyzes the data with machine learning (ML) methods to determine the AD mode attributes of CAVs. In this work, we employ the ML method, which allows for efficient and reasonable feature classification in addition to robust data fitting and function mapping capabilities [38]. The CAVs that share the same or similar driving habits are then directed to cluster together.

**Table 1: Confusion matrix of distracted driver detection database**

|      | SD   | TRH  | TPRH | TLH  | TPLH | OR   | D    | RB   | HM   | TP   |
|------|------|------|------|------|------|------|------|------|------|------|
| SD   | 91.4 | 1    | 0.8  | 0.5  | 1.1  | 0    | 0.2  | 0    | 0    | 0    |
| TRH  | 0    | 91.6 | 1.4  | 1.1  | 0.9  | 0    | 0    | 0    | 0    | 0    |
| TPRH | 2.3  | 0.7  | 90.2 | 1.6  | 0    | 0    | 0.2  | 0    | 0    | 0    |
| TLH  | 2.3  | 0    | 1.4  | 90.6 | 0.7  | 0    | 0    | 0    | 0    | 0    |
| TPLH | 0.7  | 0.2  | 0    | 1.1  | 92.8 | 0    | 0    | 0    | 0    | 0    |
| OR   | 0    | 0    | 0    | 0    | 0    | 89.2 | 0    | 0    | 1    | 1.6  |
| D    | 0    | 0    | 0    | 0    | 0    | 0    | 93.7 | 6.6  | 0.4  | 1.7  |
| RB   | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 90.6 | 2.6  | 1.9  |
| HM   | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 9.6  | 89.7 | 0.7  |
| TP   | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 3.3  | 0.7  | 89.5 |

#### 4. CONCLUSION

We provide a pipeline for anomaly detection that leverages the best features of cameras and lidar to save time and effort. We didn't try to develop a single model that would work everywhere; rather, we zeroed down on the optimal ways to mix different kinds of data. It has already been established that our threshold-based anomaly detection technique is a good foundation. The paucity of freely accessible multimodal datasets for anomaly identification was a significant challenge. Unfortunately, we were only able to provide a qualitative overview of typical cases because the dataset lacked ground truth labels for anomalies. Li et al. [62] recently introduced CODA, a real-world road corner case dataset, which includes scenes from major real-world object recognition benchmarks [63]-[65], each of which contains at least one corner case that poses a safety problem for self-driving vehicles or their surroundings. Camera and lidar data are included in every scenario. However, the entire dataset is not available at this moment. Future work in anomaly detection can be evaluated quantitatively using this dataset.

#### REFERENCES

- [1] Mercedes-Benz Group, "First internationally valid system approval for conditionally automated driving," 2021, accessed: 12.02.2022. [Online].
- [2] On-Road Automated Driving (ORAD) committee, *SAE-J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, 2021.
- [3] Torc Robotics, "Torc robotics to expand self-driving truck testing to new mexico with test center in albuquerque," 2020, accessed: 12.02.2022. [Online]. Available: <https://torc.ai/torc-robotics-to-expand-self-driving-truck-testing-to-new-mexico-with-test-center-in-albuquerque/>
- [4] Rebecca Bellan, "Tusimple completes its first driverless autonomous truck run on public roads," 2021, accessed: 12.02.2022. [Online]. Available: <https://techcrunch.com/2021/12/29/tusimple-completes-its-first-driverless-autonomous-truck-run-on-public-roads/>
- [5] Waymo, "Expanding our testing in san francisco," 2021, accessed: 12.02.2022. [Online]. Available: <https://blog.waymo.com/2021/02/expanding-our-testing-in-san-francisco.html>
- [6] G. Di Biase, H. Blum, R. Siegwart, and C. Cadena, "Pixel-wise Anomaly Detection in Complex Driving Scenes," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.

- [7] K. J. Joseph, S. H. Khan, F. S. Khan, and V. N. Balasubramanian, "Towards open world object detection," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [8] Saravanakumar, S., & Saravanan, T. (2023). Secure personal authentication in fog devices via multimodal rank-level fusion. *Concurrency and Computation: Practice and Experience*, 35(10), e7673.
- [9] D. Bogdoll, J. Breitenstein, F. Heidecker, M. Bieshaar, B. Sick, T. Fingscheidt, and M. Zoellner, "Description of Corner Cases in Automated Driving: Goals and Challenges," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2021.
- [10] K. Wong, S. Wang, M. Ren, M. Liang, and R. Urtasun, "Identifying Unknown Instances for Autonomous Driving," *arXiv:1910.11296*, 2019.
- [11] P.J. Harrison, H. Wieslander, N. Pielawski, K. Kartasalo, A. Gupta, G. Partel, L. Solorzano, A. Suveer, A.H. Klemm, O. Spjuth, et al., "Deep learning in image cytometry: a review", *Cytometry Part A*, Vol. 95(4), 366–380, 2019.
- [12] Saravanakumar, S., & Thangaraj, P. (2019). A computer aided diagnosis system for identifying Alzheimer's from MRI scan using improved Adaboost. *Journal of medical systems*, 43(3), 76.
- [13] Saravanan, T., & Saravanakumar, S. (2022). Enhancing investigations in data migration and security using sequence cover cat and cover particle swarm optimization in the fog paradigm. *International Journal of Intelligent Networks*, 3, 204-212.
- [14] X. Liu, M. Luo, P. Zhang, W. Wang, W. Huang, "video based abnormal driving behavior detection via deep learning fusion", In *IEEE Access*, Vol 7, pp. 64571- 64582., 2019.
- [15] P. Peddi, "Design of Simulators for Job Group Resource Allocation Scheduling In Grid and Cloud Computing Environments", ISSN: 2319- 8753 Vol. 6(8), pp: 17805-17811, 2017.
- [16] K. ByoungChul, K. Sooyeong, J. Mira, and N. Jae-Yeal, "Driver facial landmark detection in real driving situations", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 28(10), pp. 2753–2767, 2017
- [17] Saravanan, T., Saravanakumar, S., Rathinam, G. O. P. A. L., Narayanan, M., Poongothai, T., Patra, P. S. K., & Sengan, S. U. D. H. A. K. A. R. (2022). Malicious attack alleviation using improved time-based dimensional traffic pattern generation in uwsn. *Journal of Theoretical and Applied Information Technology*, 100(3), 682-689.
- [18] S. Aaqib, T. Stojan, K. Maurice, and E. Jan van , "Deep physiological arousal detection in a driving simulator using wearable sensors", In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 486–493, 2017.
- [19] L. Nanxiang, M. Teruhisa, and T. Fei, "Understand driver awareness through brake behavior analysis: Reactive versus intended hard brake", In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1523–1528. 2017.
- [20] L. Brun, A. Saggese, B. Cappellania, and M. Vento, "Detection of anomalous driving behaviors by unsupervised learning of graphs", In *2014 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 405–410, 2014.
- [21] T. Lossau, H. Nickisch, T. Wissel, R. Bippus, H. Schmitt, M. Morlock, and M. Grass, "Motion estimation and correction in cardiac ct angiography images using convolutional neural networks" *Computerized Medical Imaging and Graphics*, Vol. 7(6), pp. 101-110, 2019.
- [22] Saravanakumar, S. (2020). Certain analysis of authentic user behavioral and opinion pattern mining using classification techniques. *Solid State Technology*, 63(6), 9220-9234.
- [23] Q. Wan, G. Peng, Z. Li, F. Inomata, Y. Zheng and Q. Liu, "Using Asymmetric Theory to Identify Heterogeneous Drivers' Behavior Characteristics Through Traffic Oscillation," in *IEEE Access*, vol. 7, pp. 106284-1
- [24] Thangavel, S., & Selvaraj, S. (2023). Machine Learning Model and Cuckoo Search in a modular system to identify Alzheimer's disease from MRI scan images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 11(5), 1753-1761.