Proceedings of the 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS-2025)
IEEE Xplore Part Number: CFP25AZ5-ART; ISBN: 979-8-3315-1250-7

Empowering Autonomous Vehicles to Make Challenging Options in Unexpected Circumstances with Hybrid Learning

Sajeev Ram Arumugam
Department of CSE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
imsajeev@gmail.com

Sheela Gowr P
Department of CSE
Vels Institute of Science, Technology & Advanced Studies
Chennai, India
psheelagowr85@gmail.com

Ananth J P
Department of CSE
Dayananda Sagar University, Bangalore, India
jude.ananth@gmail.com

Sankar Ganesh Karuppasamy
Department of CSE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
prof.sankarganesh@gmail.com

Palani S
Department of Computer Science and Engineering (Cyber Security)
Sri Krishna College of Engineering and Technology
Coimbatore, India
s.palani87@gmail.com

Elavarasi J
Department of Artificial Intelligence and Data Science
Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College
Chennai, India
elavarasivela@gmail.com

2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) | 979-8-3315-1250-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSCDS65426.2025.11167771

***Abstract*— Autonomous vehicles (AVs) must navigate challenging and unexpected circumstances while guaranteeing security and competence. Prescribed rule-based classifications strive to handle the large unpredictability of virtual driving situations. In the proposed work, a novel hybrid architecture enables autonomous vehicles to make human-like choices in unexpected scenarios by using a combination of deep learning and data-driven planning techniques. The framework combines VOLOv7-based perception, multimodal transformers for fusing LiDAR, radar, and camera data, and a dual-policy approach using DAgger and Decision Transformer to obtain both sensitive and deliberate decision-making behaviors. An ensemble voting mechanism combines policy outputs to improve reliability. The proposed work is trained and evaluated using the Waymo Open Dataset and CARLA simulator. The proposed work attains a collision rate of 3.4%, route completion of 97.2%, and an average intervention frequency of 0.4.**

Keywords*— *Autonomous Vehicles, Deep Learning, Decision-Making, Reinforcement Learning, Imitation Learning, Human-like Reasoning

I. INTRODUCTION

Autonomous vehicles (AVs) have the potential to modernize transit by refining security, effectiveness, and user-friendliness. However, one of the most important challenges in attaining complete independence lies in allowing vehicles to make difficult, concurrent decisions in unexpected situations [1]. Current progressions in artificial intelligence, mainly deep learning, give auspicious solutions by letting AVs acquire massive amounts of driving data. By studying ideas in human behaviour, neural networks can foresee suitable answers to unique circumstances, imitating the perception of skilled drivers [2].

Techniques such as deep reinforcement learning (DRL) [3] and imitation learning (IL) [4] allow AVs to refine their decision-making policies through constant interaction with virtual and real-world surroundings. The proposed paper evaluates how deep learning can improve AV decision-making by merging understanding, forecasting, and analysis into a combined framework. A hybrid model is proposed that influences both imitation learning and reinforcement. The experiments indicate that this approach outdoes conventional methods in managing difficult situations while retaining safety and competence.

II. RELATED WORK

Study of autonomous driving has grown considerably over the former period, with various prototypes proposed to discourse the concerns of understanding, development, and management [5]. Traditional approaches depend on architectures that distinguish understanding, translation, forecasting, and decision-making. These systems normally use handmade rules and required reasoning for development and management, such as finite state mechanisms and behaviour trees [6]. While these approaches provide high clarity and consistency in comprehensible situations, they are unstable when subject to sudden or exceptional circumstances that drop excepting predefined rules.

However, end-to-end learning methods have gained traction for their potential to directly map sensor inputs to driving movements using deep neural networks. Revolutionary work introduced a convolutional neural network (CNN) [7] trained to forecast navigation angles from camera images using supervised learning. Regardless of auspicious results in precise conditions, such models often suffer from poor generality, a deficit of clarity, and incompetence to purpose about long-term moments. Furthermore, they struggle in situations where labelled data is uncommon or unclear, such as close-shave instances or uncommon road behaviours.

Addressing these contests, recent research has gradually looked at learning-based decision-making structures [8] that integrate imitation learning (IL) with reinforcement learning (RL) [9]. Imitation learning [10], mainly behaviour cloning and inverse reinforcement learning [11], allows models to study strategies by sensible expert evidence, apprehending human-like driving patterns. However, these models are subject to covariate shift, where small errors meld over time, forcing the agent into unfamiliar states. Reinforcement learning permits agents to analyze and learn optimum behaviours through trial-and-error connections with an environment. Techniques such as Deep Q-Networks (DQN) [12], Proximal Policy Optimization (PPO) [13], and Soft Actor-Critic (SAC) [14] have shown achievement in virtual driving tasks. However, pure RL methods need wide-ranging surveys and often struggle with sample efficiency and safety in real-world distribution.

Hybrid methods have recently developed as a robust plan for attaining both effective learning and potent decision-making. Methods such as DAgger (Dataset Aggregation) [15] address covariate shifts in imitation learning by iteratively gathering remedial expert data. Others modify policies using IL and improve them with RL in the model, allowing safe examination of precarious situations. In spite of this development, many existing methods still drop the ball in recreating the adaptable, environment-conscious decision-making demonstrated by human drivers. The proposed work is developed on these basics by proposing a deep learning-based framework to model human-like reasoning more efficiently. Using large-scale real-world datasets and high-integrity models, the outline aims to improve decision-making in unexpected situations pushing AV competences closer to human-level understanding.

III. Methodology

This section summarizes the proposed framework aimed at representing man-like perception in autonomous vehicles by incorporating deep neural networks (DNNs), imitation learning (IL), and reinforcement learning (RL). The methodology consists of four key stages: (1) perception through YOLOv7, (2) sensor fusion and scene appreciative is using a multimodal transformer, (3) policy learning through a combination of imitation learning and Decision Transformers, and (4) robustness improvement using ensemble methods. This architecture is intended to simplify across diverse, unforeseen situations by learning both spatial and temporal patterns from expert exhibits and simulated connections. Fig.1 shows the architecture of the proposed method.

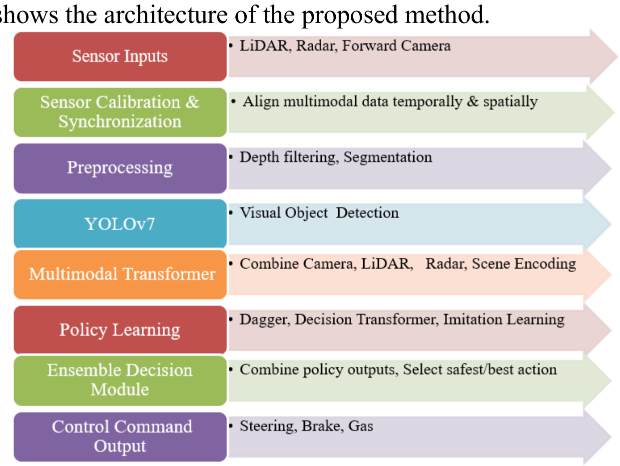

Fig.1. Multimodal Deep Learning Architecture

A. Data Collection and Preprocessing

The Waymo Open Dataset [16] serves as a real-world resource, contributing corresponding multi-sensor data (cameras, LiDAR, Radar), useful observations, and skilled driving routes across various civic environments. For creating hazardous driving circumstances that are durable to detain in real life, the CARLA simulator allows full control over conservational variables, traffic representatives, and sensor formations, making it perfect for collecting cooperating data using algorithms like Dagger [17] and for training Decision Transformers in long-horizon tasks. Proficient policies are resultant either from human drivers in the Waymo dataset or scripted agents in CARLA [18], providing ground truth actions for imitation learning. Observations, including object detection labels, routes, and control signals, are produced automatically in replication and used directly or improved through manual labelling tools in real-world data. This hybrid data collection approach guarantees complete attention both routine and rare situations, allowing vigorous training of observation, development, and decision modules.

Preprocessing LiDAR data starts with sensor calibration and synchronization, where basic and outermost parameters are used to align data from multiple sensors into one frame. This process confirms spatial and temporal reliability across diverse modalities. Depth filtering comprises removing noise and outliers from the raw LiDAR point clouds using methods like statistical outlier removal, voxel grid down sampling, and range-based filtering to retain only significant spatial information. Data segmentation is applied to separate relevant structures from the scene, such as roads, vehicles, pedestrians, and buildings. These preprocessing steps convert raw, unstructured sensor data into structured inputs appropriate for combination in multimodal transformers, object detection networks like YOLOv7 [19], and decision-making frameworks such as DAgger and Decision Transformers.

B. Perception Module

The first phase of the pipeline is a real-time perception module based on YOLOv7, an advanced object detection network. It processes the frontal camera to identify both dynamic and static entities such as automobiles, walkers, bicycles, traffic signs, and lane marks. The output is a set of bounding boxes $B = \{b_i\}_{i=1}^{N}$, each with a class label ci, and confidence score si. The detections are transformed into a structured representation:

$$O_t = \{(c_i, s_i, x_i, y_i, w_i, h_i)\}_{i=1}^{N} \qquad (1)$$

Where (x_i, y_i) is the center, and (w_i, h_i) are the height and weight of the bounding box respectively. This data helps as a visual input for the fusion module and acts as the basis for situational consciousness.

C. Multimodal Transformer

Autonomous vehicles depend on various sensory inputs, including visualization, LiDAR, radar, and GPS. A multimodal transformer is used to combine these diverse signals into an incorporated, high-dimensional scene demonstration. Let *It* is the image features extracted from YOLOv7; *Lt* is the LiDAR point cloud features; *Mt* is the HD map and lane topology; and St is the ego vehicle's state vector (velocity, acceleration, heading). The inputs are determined into embedding vectors:

$$E_t = TransformerEncoder(concat(I_t, L_t, M_t, S_t)) \qquad (2)$$

The transformer utilizes self-attention to evaluate the significance of each modality vigorously.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (3)$$

The result provides the model to acquire a rich spatial and semantic framework crucial for downstream decision-making.

D. Policy Learning

Imitation Learning Initialization

The behaviour cloning is used to bootstrap the policy, where a policy $\Pi\boldsymbol{\theta}$ is trained using supervised learning on proficient routes

$$\pi_\theta(a_t|s_t) \approx a_t^{expert} \qquad (4)$$

Minimizing the loss:

$$L_{BC}(\theta) = E_{(s_t,a_t)}{\sim}D_{expert}[\|\pi_\theta(s_t) - a_t\|^2] \qquad (5)$$

DAgger for Expert Correction

To address covariate shift, the DAgger (Dataset Aggregation) algorithm is employed. In DAgger, the policy relates to the situation, and the proficient offer correct labels for visited states, which are combined into the dataset.

$$D \leftarrow D \cup \left\{\left(s_t, a_t^{expert}\right)\right\}_{t=1}^{T} \qquad (6)$$

The policy is iteratively restructured on the combined dataset to enhance generalization under the distribution persuaded by its movements.

Decision Transformer

To improve decision-making over prolonged sequences, the Decision Transformer (DT) is used, which frames reinforcement learning as a sequence modeling problem. The input is a sequence of *Rt* as returns-to-go, s_t as observed states, a_{t-1} as Past actions. The Decision Transformer calculates the next act using an autoregressive transformer:

$$a_t = fDT(R_t, s_1, a_1, \ldots . s_{t-1}, a_{t-1}, s_t) \qquad (7)$$

The above expression (7) permits the model to focus on chosen outcomes and reason about the multi-step effects of actions, allowing high-level strategic decisions even in innovative circumstances.

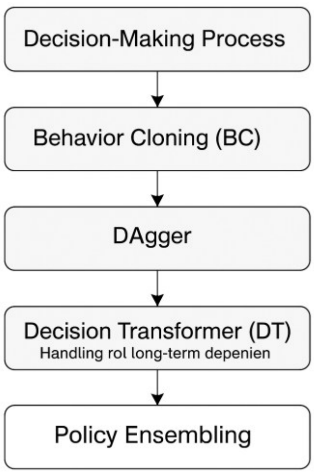

Fig 2. Decision pipeline

The decision-making process begins with Behavior Cloning to initialize the policy using supervised learning. Once the policy is trained, DAgger is employed, which iteratively collects new data where the model updates. This allows a transient shift and improves robustness. To handle longer-term dependencies the Decision Transformer is used which takes decision-making as a sequence modeling task using a transformer architecture. Finally, multiple policies from BC, DAgger, and DT are combined to form a robust output, reducing variance and improving safety in ambiguous or edge-case scenarios. The ensemble votes or averages predictions based on uncertainty estimates to choose the final action. The overall Process is visually represented in Fig. 2.

E. Ensemble Models

To further enhance reliability, model ensembling is used, where multiple autonomously trained policy networks vote or average their predictions:

$$\widehat{a_t} = \frac{1}{K}\sum_{k=1}^{K} \pi_{\theta k}(s_t) \qquad (8)$$

Instead, ambiguity can be captured using the inconsistency of outputs:

$$\sigma_t^2 = \frac{1}{K}\sum_{k=1}^{K}(\pi_{\theta k}(s_t) - \widehat{a_t})^2 \qquad (9)$$

In confusing situations, the vehicle can choose traditional emergency behaviors such as stopping or complying with rule-based control.

IV. Experimental Setup

The experimental validation of the proposed autonomous vehicle decision-making system was shown through a complete, multi-stage testing protocol that is considered to calculate performance through both simulated and real-world surroundings. For real-world data, the Waymo Open Dataset is used, which offers extensive multi-modal sensor data, including high-resolution LiDAR, radar, and multiple camera views. The proposed work used 5000 hours of data. This dataset is analytical for preparing the perception models and understanding expert routes in real driving situations. The experimental framework incorporates several key components, comprising simulation environments, hardware configurations, software employments, training methods, and evaluation metrics, each cautiously considered to validate the system's proficiencies. For virtual reality testing, CARLA 0.9.14 [20] is an open-source autonomous driving simulator that offers high-fidelity built-up surroundings with active weather and lighting situations. The robust simulator's sensor models enable the creation of various testing situations, starting from routine driving circumstances to complex edge cases. The simulation surroundings were further enriched with variable weather patterns, including rain, fog, and nighttime conditions, to calculate the system's performance under hostile conditions.

The hardware configuration was selected to support both training and real-world deployment requirements. For training, workplaces enlightened with four NVIDIA A100 80GB GPUs, AMD EPYC 7763 processors with 64 cores, 512GB of DDR4 RAM, and a 20TB NVMe SSD array is used for effective data management. The software pile was created to incorporate the various mechanisms of the autonomous driving system. The perception module exploited VOLOv7 implemented in PyTorch 2.1 for object detection, together with a Multi-Modal Transformer enhanced through TensorRT 8.6 for efficient sensor fusion. The decision-making subsystem used a Decision Transformer structure made on JAX 0.4.13, while DAgger used a custom Python hybrid architecture for optimum execution.

The dataset is divided into three subsets: training (70%), validation (15%), and testing (15%). The training set contains a combination of standard driving scenarios, highway scenarios, and initial edge cases created through scripted agents. The validation set is used for hyperparameter tuning and to prevent overfitting during training, while the test set includes unobserved and more difficult edge-case situations to assess the ability of the proposed framework. To increase diversity, each split sustains stability across diverse driving situations, movement densities, and weather circumstances. The experimental results showed the system's performance in managing difficult driving situations compared to standard approaches, mainly in edge cases involving unexpected walker behavior and unfavorable weather conditions. The testing procedure verified the effectiveness of the distinct technical components and established their unified incorporation into an autonomous driving system capable of making human-made decisions in complex situations.

V. Results and Discussions

Comparative analysis is a vital factor of our experimental endorsement. The proposed system is estimated against three baseline architectures: a rule-based FSM [21], a CNN approach based on NVIDIA's PilotNet architecture [22], and a modular pipeline, Apollo 7.0 [23] for decision-making models. These comparisons are used to measure performance enhancements. Fig. 3 shows the performance metrics in terms of success rate, intervention rate, and planning accuracy.

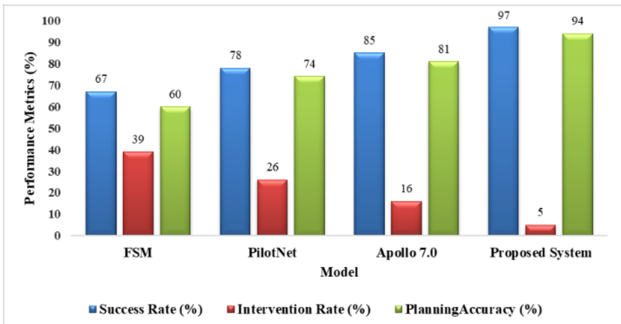

Fig.3. Performance Metrics for decision making

The comparative analysis graph of collision rate, route completion, and average intervention frequency for the FSM, PilotNet, Apollo 7.0, for driving performance and the proposed system is shown in Fig. 4. The Collision Rate calculates the amount of test incidents in which the autonomous vehicle accounts for a collision. The Route Completion metric signifies the percentage of the scheduled route that the autonomous vehicle effectively finishes, deprived of major letdowns. Average Intervention Frequency calculates how often a human proficient must intrude or take control during an incident due to security or performance matters.

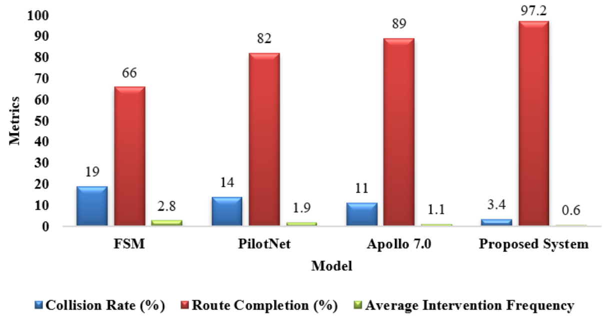

Fig.4 Performance Metrics for driving performance

The comparative analysis graph, Fig. 5, shows normalized scores for data efficiency, edge case handling, and compute cost across Pure Imitation, Pure RL, and the proposed work. Data Efficiency mentions how efficiently a learning algorithm can excerpt valuable patterns or behaviors from the training data. Edge Case Handling assesses how well a model can manage unusual, risky situations. Compute Cost represents the computational resources needed to train and use the model.

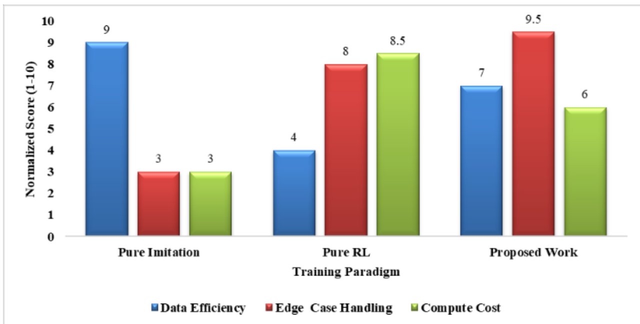

Fig.5 Training Paradigm Comparisons

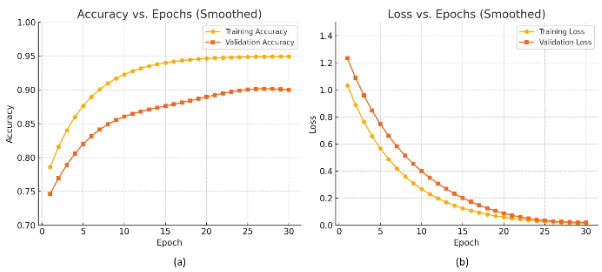

Fig.6. Training and validation performance: (a)Accuracy vs. Epochs, (b)Loss vs. Epochs.

As shown in Fig. 6, the training loss steadily decreased, converging after around 25 epochs with minimal overfitting. The validation accuracy spiked at 91%, while training accuracy reached 95%, indicating good generalization. The use of DAgger corrections helped reduce validation loss spikes, which are typically caused by covariate shifts.

VI. Conclusion

The proposed work presents a robust and adaptive decision-making framework for autonomous vehicles that incorporates VOLOv7-based perception, multimodal transformers for context fusion, and a hybrid policy architecture combining DAgger, Decision Transformer, and ensemble learning. Complete wide sets of experiments are done using both the Waymo Open Dataset and CARLA simulations. The proposed system proves higher performance across key metrics such as collision rate, route completion, and intervention frequency, considerably outclassing conventional rule-based and deep learning baselines. By imitating human-like perception and enabling an overview of unseen situations, the proposed approach provides an auspicious step toward safety, consistency, and smart independence in difficult real-world situations.

References

[1] D. Parekh et al., "A review on autonomous vehicles: Progress, methods and challenges," *Electronics*, vol. 11, no. 14, pp. 2162, 2022.

[2] K. Wang, C. Shen, X. Li, and J. Lu, "Uncertainty Quantification for Safe and Reliable Autonomous Vehicles: A Review of Methods and Applications," *IEEE Trans. on Intelligent Transportation Systems*, vol. 26, no. 2, pp. 1123–1145, 2025.

[3] N. Botteghi, M. Poel, and C. Brune, "Unsupervised representation learning in deep reinforcement learning: A review," *IEEE Control Systems Magazine*, vol. 45, no. 2, pp. 26–68, 2025.

[4] M. Zare, P. M. Kebria, A. Khosravi, and S. Nahavandi, "A survey of imitation learning: Algorithms, recent developments, and challenges," *IEEE Trans. on Cybernetics*, vol. 54, no. 1, pp. 112–130, 2024.

[5] A. D. M. Ibrahum, M. Hussain, and J.-E. Hong, "Deep learning adversarial attacks and defenses in autonomous vehicles: A systematic literature review from a safety perspective," *Artif. Intell. Rev.*, vol. 58, no. 1, pp. 1–53, 2025.

[6] S. Chen, X. Hu, J. Zhao, R. Wang, and M. Qiao, "A review of decision-making and planning for autonomous vehicles in intersection environments," *World Electric Vehicle Journal*, vol. 15, no. 3, p. 99, 2024.

[7] G. Doğan and B. Ergen, "A new CNN-based semantic object segmentation for autonomous vehicles in urban traffic scenes," *Int. J. of Multimedia Information Retrieval*, vol. 13, no. 1, pp. 11–22, 2024.

[8] B. Wang, X. Gong, P. Lyu, and S. Liang, "Iterative learning-based cooperative motion planning and decision-making for connected and autonomous vehicles coordination at on-ramps," *IEEE Trans. on Intelligent Transportation Systems*, vol. 25, no. 3, pp. 2098–2110, 2024.

[9] J.-H. Choi, D. Kim, J.-S. Yoo, B.-J. Kim, and J.-T. Hwang, "Enhancing Autonomous Driving with Pre-trained Imitation and Reinforcement Learning," in *Proc. 2025 Int. Conf. on Electronics, Information, and Communication (ICEIC)*, Jeju Island, South Korea, 2025, pp. 1–3.

[10] N. A. Ajak, W. H. Ong, and O. A. Malik, "A Comparison of Imitation Learning Pipelines for Autonomous Driving on the Effect of Change in Ego-vehicle," in *Proc. 2024 IEEE Intelligent Vehicles Symposium (IV)*, Seoul, South Korea, 2024, pp. 1693–1698.

[11] S. Deshpande, R. Walambe, K. Kotecha, G. Selvachandran, and A. Abraham, "Advances and applications in inverse reinforcement learning: a comprehensive review," *Neural Computing and Applications*, vol. 36, no. 1, pp. 1–53, 2025.

[12] K. Vinoth and P. Sasikumar, "Multi-sensor fusion and segmentation for autonomous vehicle multi-object tracking using deep Q networks," *Scientific Reports*, vol. 14, no. 1, p. 31130, 2024.

[13] M. Bilban and O. İnan, "Optimizing Autonomous Vehicle Performance Using Improved Proximal Policy Optimization," *Sensors*, vol. 25, no. 6, p. 1941, 2025.

[14] Y. Chen and S. Wu, "Framework of active obstacle avoidance for autonomous vehicle based on hybrid soft actor-critic algorithm," *J. Transp. Eng. Part A: Systems*, vol. 149, no. 4, p. 4023002, 2023.

[15] J. Ahn, S. Shin, J. Koo, M. Kim, and J. Park, "Data Aggregation (DAgger) Algorithm Using Adversarial Agent Policy for Dynamic Situations," in *Proc. Int. Conf. on Intelligent Autonomous Systems*, Cham, Switzerland: Springer, 2023, pp. 91–104.

[16] Waymo, "About – Waymo Open Dataset," [Online]. Available: https://waymo.com/open/. Accessed: Apr. 10, 2025.

[17] J. Wellander and A. Petersén, "Autonomous Driving via Imitation Learning in a Small-Scale Automotive Platform," *Chalmers Open Digital Repository*, 2024.

[18] L. Vezzani, F. Strada, and A. Bottino, "Towards CARLA VR: Testing Head-Up Displays in CARLA Driving Simulator," in *Future of Information and Communication Conference*, Cham, Switzerland: Springer, 2025, pp. 1–16.

[19] S. Mammeri, M. Amroune, M.-Y. Haouam, I. Bendib, and A. Corrêa Silva, "Early detection and diagnosis of lung cancer using YOLO v7, and transfer learning," *Multimedia Tools and Applications*, vol. 83, no. 10, pp. 30965–30980, 2024.

[20] Qayyum, Adnan, Muhammad Usama, Junaid Qadir, and Ala Al-Fuqaha. "Securing connected & autonomous vehicles: Challenges posed by adversarial machine learning and the way forward." *IEEE Communications Surveys & Tutorials* 22, no. 2 (2020): 998-1026.

[21] S. R. Arumugam, B. Ravichandran, D. Baskaran, and R. Annamalai, “Lung Lobe Segmentation and Lung Cancer Detection with Hybrid Optimization-Enabled Deep Learning Using CT Images,” *Journal of Mechanics in Medicine and Biology*, World Scientific, Nov. 2024, doi: 10.1142/S0219519424500477.

[22] D. Muralitharan, S. Tamilselvi, J. D. Joann, S. R. Arumugam, S. G. Karuppasamy and C. S. Shylaja, "Exploring the Efficacy of U-Net-CNN Architecture for Automated Detection of Psoriatic Arthritis from Medical Imaging Data," *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, Erode, India, 2024, pp. 227-232, doi: 10.1109/ICSSAS64001.2024.10760526..

[23] H.-Y. Jung, D.-H. Paek, and S.-H. Kong, "Open-Source Autonomous Driving Software Platforms: Comparison of Autoware and Apollo," *arXiv preprint arXiv:2501.18942*, 2025.