Deep Reinforcement Learning for Autonomous Vehicles

Srinivasa Kalyan Vangibhurathachhi

Srinivasa.Kalyan2627@gmail.com

Abstract
This study investigates the application of Deep Reinforcement Learning (DRL) in enhancing the performance and decision-making capabilities of Autonomous Vehicles (AVs). By leveraging DRL techniques such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Asynchronous Advantage Actor-Critic (A3C), AVs can learn to make intelligent, real-time decisions in dynamic environments. The paper uses a comparative analysis of these DRL algorithms to evaluate their effectiveness in path planning, lane-keeping, and obstacle avoidance tasks. Key challenges encountered in deploying DRL in AVs include scalability, safety, real-time decision-making, and the sim-to-real transfer of models. The study also highlights the role of simulation platforms like CARLA and OpenAI Gym in training DRL models and discusses their impact on model reliability and real-world performance. The findings suggest that while DRL shows great promise for improving AV capabilities, challenges remain in its practical application, particularly regarding the safety and ethical implications of decision-making in critical driving situations. Future research directions include enhancing simulation environments, reward function design, and developing robust safety protocols to ensure the safe deployment of DRL-powered AVs in real-world scenarios.

Keywords: Deep Reinforcement Learning, Autonomous Vehicles, Path Planning, Drl Algorithms, Simulation Environments

1. Introduction

1.1 Background on Autonomous Vehicles

Autonomous vehicles (AVs) are transforming transportation, promising improvements in road safety, traffic efficiency, and mobility. Operating without human intervention, AVs use sensors, machine learning, and artificial intelligence (AI) to navigate their environment (Soori et al., 2023). The development of AV technology has been rapidly advancing, driven by major companies like Tesla, Google (Waymo), and Uber. AVs have the potential to significantly reduce traffic accidents caused by human error, which accounts for over 90% of incidents (Soori et al., 2023). They can also reduce traffic congestion, improve energy efficiency, and increase mobility for those unable to drive, such as the elderly or disabled. However, challenges remain, particularly in ensuring AV safety in dynamic, unpredictable environments. AVs must make real-time decisions based on numerous factors, including traffic conditions and human behaviors (Xie et al., 2025). Moreover, integrating AVs into existing road systems poses regulatory, ethical, and infrastructural challenges. This is where AI, particularly reinforcement learning, becomes critical in AV development.

Deep Reinforcement Learning (DRL) is a machine learning method that combines deep learning with reinforcement learning. DRL enables an agent to learn optimal decision-making through trial and error, using feedback from its environment (Chen, 2022). In AVs, DRL allows the vehicle to improve its driving policies over time by interacting with various driving conditions (Zhu & Zhao, 2021). It is particularly advantageous in complex environments, as it can adapt and refine its decision-making without explicit human programming. Unlike traditional rule-based systems, DRL can adjust to unforeseen scenarios, making it highly suitable for real-time decision-making in driving (Tian et al., 2025). DRL's importance in AVs lies in its ability to handle dynamic environments and enable real-time responses to changing traffic conditions (Antonio & Maria-Dolores, 2022). While traditional methods struggle in such environments, DRL allows AVs to continuously learn and improve from their experiences, making it a more scalable and adaptable approach to autonomous driving. DRL’s ability to improve over time without requiring manual updates positions it as a cornerstone for AV development, enabling them to navigate the unpredictable nature of real-world driving.

1.2 Problem statement

Autonomous vehicle (AV) control faces significant challenges due to the complexity of real-world environments. AVs must navigate dynamic conditions, such as varying traffic patterns, unpredictable behaviour from pedestrians, and interactions with other vehicles (Rezwana & Lownes, 2024). Additionally, diverse road users, such as cyclists and motorcyclists, further complicate decision-making (Useche et al., 2025). These complexities require AVs to make split-second decisions in a constantly changing environment, making it crucial for their control systems to respond effectively and in real-time.

Traditional approaches, such as rule-based systems and supervised learning, fall short in these dynamic environments. Rule-based systems are limited by predefined actions and fail to account for the unpredictability of real-world scenarios (Varshney & Torra, 2023). Supervised learning systems require large amounts of labelled data and often struggle to generalise to new situations, leaving gaps in the AV’s ability to handle edge cases (Sarker, 2021). These methods lack the flexibility and adaptability needed for real-time, complex decision-making. Deep Reinforcement Learning (DRL) offers a solution by enabling AVs to learn and adapt based on their experiences continuously. DRL allows AVs to improve decision-making over time, optimise their responses to real-world challenges, and ensure they can navigate complex environments more efficiently and safely.

1.3 Research Objectives and Contributions

This paper explores the use of Deep Reinforcement Learning in autonomous vehicles, aiming to:

- Review the state-of-the-art DRL algorithms applied to autonomous driving tasks.
- Identify the key challenges in applying DRL to AVs and suggest potential solutions.
- Discuss the real-world implications of using DRL in AVs, including safety, ethical concerns, and regulatory challenges.

This paper contributes to the growing body of knowledge on intelligent systems for transportation by synthesizing current research and offering new insights into the application of DRL in autonomous driving.

2. Literature Review

2.1 Existing Techniques in Autonomous Driving

Haque et al. (2022) observe that early autonomous driving systems relied on rule-based approaches and supervised learning models. Varshney and Torra (2023) state that rule-based systems function by implementing predefined instructions for specific tasks, such as lane-keeping and obstacle avoidance. However, these systems struggled with real-world variability and could not adapt to unforeseen circumstances, making them less reliable in dynamic environments (Bellone et al., 2021). Supervised learning approaches also faced limitations, as they required vast amounts of labeled data for training, which could not cover all possible driving scenarios (Chib & Singh, 2023). Furthermore, these models often struggled with generalization, meaning they would fail when presented with novel, real-world situations (Bellone et al., 2021). As the complexity of driving scenarios grew, these early systems became inadequate, leading to the exploration of more flexible learning methods such as Deep Reinforcement Learning (DRL), which provides greater adaptability and scalability.

2.2 State-of-the-Art DRL for Autonomous Vehicles

Recent advancements in autonomous driving have seen the integration of Deep Reinforcement Learning (DRL) algorithms, such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Asynchronous Advantage Actor-Critic (A3C) (Kiran et al., 2021). These DRL algorithms have proven to be effective in enabling AVs to navigate complex and dynamic environments. DQN, for instance, uses a deep neural network to approximate action-value functions and has been successfully applied in tasks like lane-changing maneuvers (Guo & Harmati, 2024). The algorithm works by associating rewards with specific actions and refining the model through iterative trials. PPO, an on-policy method, focuses on policy optimization by ensuring stability through a clipped objective function (Markowitz & Staley, 2023). This makes it well-suited for continuous decision-making tasks in urban environments. A3C, using parallel agent-based learning, enables quicker training through multi-agent systems and is especially useful in multi-agent environments like traffic intersections (Hua et al., 2023). These algorithms have already demonstrated real-world applications, with DQN being applied in lane-keeping tasks, PPO for urban driving scenarios, and A3C for resource allocation in intersections. They have outperformed earlier methods in terms of adaptability and real-time decision-making, demonstrating DRL's potential for improving AV control systems.

2.3 Challenges in DRL for Autonomous Vehicles

Real-world applications of DRL experience multiple difficulties because implementing the method in autonomous driving systems proves challenging. The main issue regarding DRL implementation involves its data requirements and computational power needs that prove problematic for real-time deployment (Kiran et al., 2021). The safety of DRL agents depends heavily on proper rewards function design because inadequate planning might result in unsafe behavioral learning. An example of such a system would be a framework allowing increased efficiency at the cost of reduced safety when making driving decisions. Another hurdle is real-time decision-making. DRL models demonstrate slow convergence during training in complex scenarios thus extending the delay in which the AV can generate fast critical driving actions (Zhu & Zhao, 2021). The inability for simulation-trained models to function optimally in genuine operational settings arises from their incapability to handle real-life conditions which include weather elements together with variabilities in road foundations and unpredictable human movements (Zhu & Zhao, 2021).

The application of handle sim-to-real situations poses difficulties which researchers are continuously trying to overcome. The differences between simulated and real-world settings cause problems when trying to use DRL-based models in uncontrolled dynamic situations unless major retraining occurs. The implementation of Deep Reinforcement Learning (DRL) in autonomous vehicles faces crucial obstacles presented in Figure 1 that combine resource requirements with data needs under safety issues and gradual real-time decision-making speed while also requiring adaptation from simulated to actual conditions and complex environment translation. DRL implementation in AV requires successful resolution of these critical obstacles.

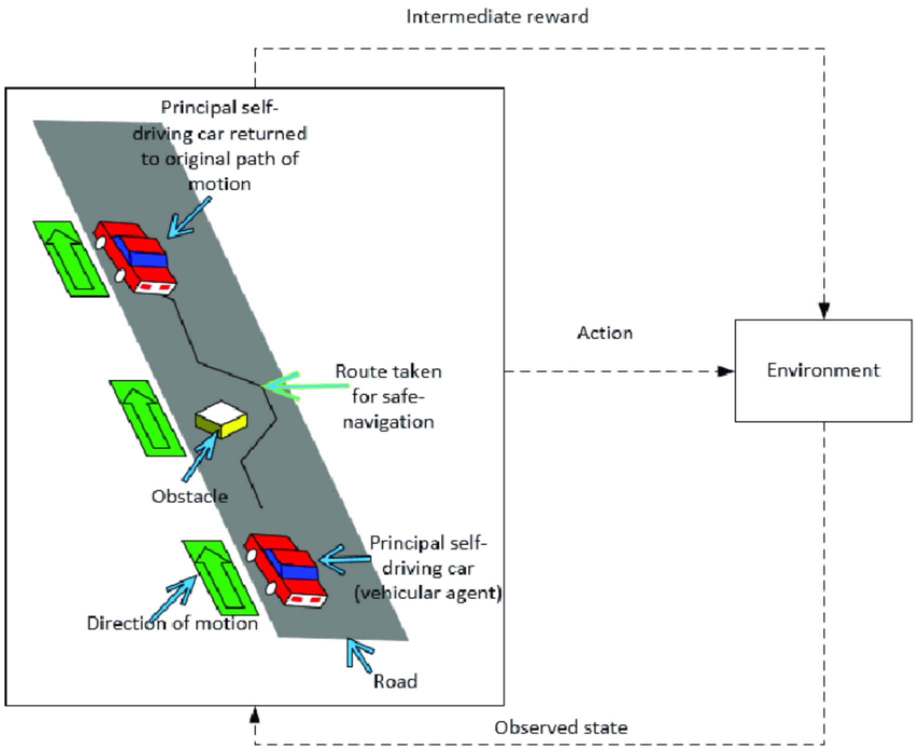

Figure 1: Deep reinforcement learning (DRL) for self-driving cars (Source: Gupta et al., 2020)

3. Proposed Solutions and Approaches

3.1 DRL Algorithms for Autonomous Vehicles

Deep Reinforcement Learning (DRL) algorithms are critical for training autonomous vehicles (AVs) to navigate complex environments and make real-time decisions (Bondre et al., 2024). Among the most commonly used DRL algorithms are Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Asynchronous Advantage Actor-Critic (A3C). DQN, introduced by Mnih and colleagues in 2015, is effective in tasks requiring discrete decisions, such as lane-keeping and path following (Kiran et al., 2021). DQN uses a deep neural network to approximate action-value functions, which allows the AV to learn optimal driving strategies by trial and error in a simulated environment (Tammewar et al., 2023).

The on-policy method PPO presented by Schulman et al. (2017) maintains stability and reliability in continuous action spaces through which it controls the acceleration and steering actions of AV control (Guo et al., 2024). The method optimizes the policy through direct updates, which guarantee safe behavior changes. Mnih et al. (2016) established A3C, which deploys multiple agents simultaneously to explore the environment while speeding up the learning process according to Liu et al. (2024). A3C brings significant benefits for managing environments containing multiple agents, including complex traffic situations with interacting vehicles, because it enhances decision-making speed in constantly changing conditions (Hua et al., 2023). Such DRL algorithms provide vehicles with capabilities to generate better driving decisions through efficient responses to unpredictable and complicated road situations. Table 1 illustrates structural and operational differences between DQN and A3C methods by using the comparison table format. The distinctions between DQN and A3C contribute to understanding different approaches within the DRL algorithms used for AVs even though PPO is not explicitly mentioned.

Test Rounds	Steps per Game	Maximum Reward		Running Time	
		DQN	A3C	DQN	A3C
100	600	2325	1735	135.6 s	133.6 s
200	600	3806	3525	284.9 s	237.3 s
400	600	2678	2565	608.8 s	455.0 s
100	1200	14,152	8246	277.2 s	245.4
200	1200	14,442	9237	617.0 s	459.4 s
100	2000	28,790	22,364	498.8 s	409.6 s

Table 1: Comparison of DQN with A3C algorithm (Source: Bi et al., 2023)

3.2 Environment Modelling (140 words)

The training of Autonomous Vehicle agents depends on simulation platforms CARLA as well as OpenAI Gym because these platforms supply protected test environments that let DRL models train without endangering actual driving experience. Providing urban-specific high fidelity simulations is CARLA while it delivers realistic conditions of roads and weather patterns as well as traffic systems (Malik et al., 2022). The simulation software lets autonomous vehicles practice advanced driving maneuvers which include lane-switching alongside parking functions and human-vehicle encounters (Zhang et al., 2024). The agents learn diverse scenarios through DRL modeling that enables safe deployment before operational use. Reinforcement learning algorithms get tested through Openai Gym which provides a general platform despite not having built-in auto vehicle capabilities (Towers et al., 2024). The platform provides simulation environments which enable AV decision-making strategy assessments in basic circumstances ahead of their implementation within complex simulation systems.

The AV training process is boosted by realistic simulation environments such as CARLA and OpenAI Gym because these systems create safe environments for testing at scale across various situations (Zhang et al., 2024). The CARLA simulator's system design is depicted in Figure 2 which illustrates how components between the virtual environment and agent along with the control system operate. Through this architecture the agent connects its sensors and actuators to ensure dynamic training situations within the simulation environment. Environmental development and testing of Deep Reinforcement Learning (DRL) models relies on this basic architectural structure to operate in a safe controlled space without risks.

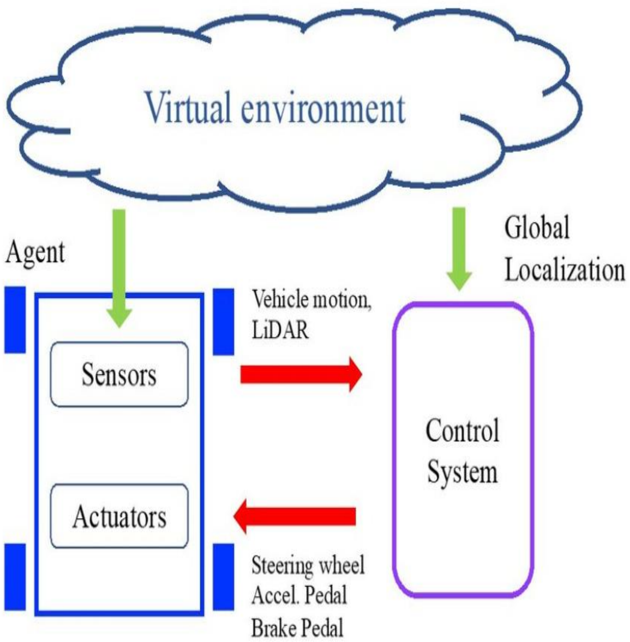

Figure 2: CARLA architecture for autonomous vehicle simulation (Terapaptommakol et al., 2022)

3.3 Reward Design

Constructing suitable reward functions stands essential because they drive DRL agents to execute appropriate behaviors. The reward system in autonomous vehicles requires equilibrium between safety elements and operational effectiveness as well as passenger comfort to guide proper vehicle decision-making (Chen et al., 2023). Inspection rewards of automated vehicles depend on penalizing dangerous behavior as well as encouraging actions that prevent collisions according to Abouelazm et al. (2024). An AV system receives improved rewards when it maintains proper distance between vehicles along with respecting traffic standards. The reward system under efficiency gives benefits for minimizing fuel usage and establishing normal driving speed and these incentives apply more to highway and lengthy distances (Zhang et al., 2023). The system gives rewards for gentle vehicle speed transitions which leads to decreased vehicle breakdowns (Karacalı et al., 2023). AVs receive rewards through their system when they steer smoothly between lanes since passenger comfort factors into the evaluation.

Through reward systems AVs enhance their operations progressively as they receive environmental feedback to optimize their decision processes. The reward functions utilized in CARLA enable autonomous vehicles to gain rewards through effective route navigation as they prevent crashes and respect traffic speed limits (Chen et al., 2024). The structural method to create reward functions for autonomous vehicles is illustrated in Figure 3. The system establishes methods to evaluate driving performance including environmental variables in order to determine reward or penalty scores. Through the logic diagram users can comprehend how safety elements along with efficiency and comfort are measured to form the reward system which drives the DRL agent to discover optimal driving protocols.

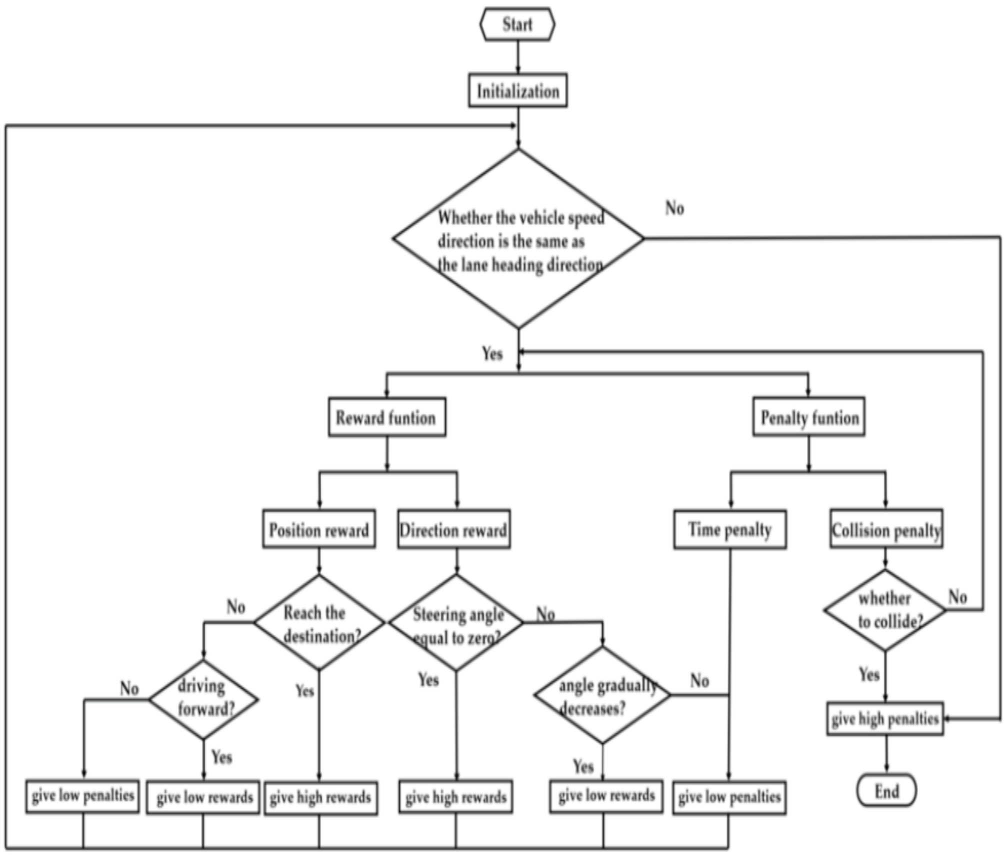

Figure 3: Logic diagram of reward function (Source; Lin et al., 2022)

4. Impact and Use Cases

4.1 Real-World Applications

Deep Reinforcement Learning (DRL) has revolutionised autonomous vehicles (AVS), enhancing their ability to navigate complex environments and make decisions in real-time. Path planning benefits significantly from DRL applications because AVs can use DRL algorithms to identify optimal routes through the evaluation of present traffic situations and street obstacles and possible threats (Reda et al., 2024). Through DRL AVs achieve enhanced navigation policy development by processing their past experiences according to Mackay et al. (2022). The application of DRL extends to impede maintenance operations in vehicles. AVs require the ability to navigate within their designated lanes as they handle environmental elements including additional traffic vehicles and road curve conditions (Guo et al., 2024). The deployment of DRL systems allows AVS to develop proper steering behaviors that efficiently prevent drift while creating secure lane changes (Lv et al., 2022).

Obstacle avoidance stands as the most fundamental element of autonomous driving because unpredictability exists in dynamic road environments (Aizat et al., 2023). Through DRL algorithms autonomous vehicles become capable of detecting pedestrians alongside other vehicles together with unexpected obstacles so they can perform real-time accident prevention decisions (Chen et al., 2024). Cars operating in urban zones and driving on highways both require these essential tasks because city roads present congested traffic and numerous pedestrian intersections and highway roads maintain slower speeds yet necessitate critical yet simple decision-making processes. DRL demonstrates effective performance in practical applications for automated vehicles according to real-world data. The Full Self-Driving (FSD) system from Tesla makes use of DRL technology for urban and highway navigation through which it both shapes optimal routes and detects intersections (Hu et al., 2025). DRL plays a vital role at Waymo and Uber ATG to enhance both urban path planning and vehicle coordination for improved automation in AV operations.

4.2 Safety, Efficiency, and Ethical Considerations

The inclusion of DRL within autonomous vehicles triggered substantial improvements of their performance capabilities alongside safety enhancements. DRL-based automated vehicles learn from their environment to handle immediate changes along with unknown events like unexpected braking or pedestrian crossings as described in Chen et al. (2024). The fast capability for making well-informed choices through DRL reduces the occurrence of accidents. Current research indicates that AVs implement DRL technology surpass traditional cars because they offer more efficient collision prevention and emergency emergency braking capabilities (Muzahid et al., 2022). The efficiency of driving behavior improves through DRL algorithms because they help minimize fuel expenses and enhance traffic management. The DRL-trained AVs modify their speed to achieve maximum efficiency while preventing traffic congestion while reducing fuel usage during extended journeys according to Du et al. (2022). The process of decision-making brings up moral questions specifically in situations where clear solutions are not obvious. The DRL system encounters ethical quandaries when forced to pick between committing two negative actions with the example of preventing one pedestrian while endangering another pedestrian (Everett et al., 2021). The decision-making challenges presented by "trolley problems" create substantial ethical issues for AV systems since they conflict with determining how algorithms should handle lethal circumstances (Poszler et al., 2023).

5. Conclusion

DRL has transformed the development of AVs through its enhancement of vehicle capabilities for conducting sophisticated computations during live operations. DRL controls autonomous vehicles through DQN, PPO, and A3C algorithms, which enable vehicles to change in dynamic environments and enhance their abilities for path planning and obstacle avoidance while keeping steady on lanes. By improving itself through environmental interactions, AVS achieves both safety and operational efficiency for new and unpredictable real-life situations, thus paving the way toward future transportation transformation. Many obstacles hinder the implementation of DRL in AVS because of problems with scalability, alongside real-time choice generation and effective transitions to actual highway environments. The proposed improvements to simulation environments and reward functions have shown progress but require additional development. Critical ethical matters need attention when making decisions about autonomous vehicle behaviors in complicated moral situations.

Additional research needs to be conducted to solve the unresolved issues that DRL presents for AV operations. Future development of DRL requires improvements in its scalability alongside enhancements in simulation-to-reality translation and ethical frameworks that determine decision processes. AV transportation systems improved through DRL integration will create safer ways to transport people that are faster and easier to access by all members of society.

References

[1] A. Abouelazm, J. Michel, and J. M. Zöllner, "A review of reward functions for reinforcement learning in the context of autonomous driving," in *2024 IEEE Intelligent Vehicles Symposium (IV)*, 2024, pp. 156–163.

[2] A. Gupta, A. S. Khwaja, A. Anpalagan, L. Guan, and B. Venkatesh, "Policy-gradient and actor-critic based state representation learning for safe driving of autonomous vehicles,"*Sensors*, vol. 20, no. 21, Art. no. 5991, 2020.

[3] A. J. M. Muzahid, S. F. Kamarulzaman, M. A. Rahman, and A. H. Alenezi, "Deep reinforcement learning-based driving strategy for avoidance of chain collisions and its safety efficiency analysis in autonomous vehicles,"*IEEE Access*, vol. 10, pp. 43303–43319, 2022.

[4] A. K. Mackay, L. Riazuelo, and L. Montano, "RL-DOVS: Reinforcement learning for autonomous robot navigation in dynamic environments,"*Sensors*, vol. 22, no. 10, p. 3847, 2022.

[5] A. K. Varshney and V. Torra, "Literature review of the recent trends and applications in various fuzzy rule-based systems,"*International Journal of Fuzzy Systems*, vol. 25, no. 6, pp. 2163–2186, 2023.

[6] A. Tammewar, N. Chaudhari, B. Saini, D. Venkatesh, G. Dharahas, D. Vora, and S. Alfarhood, "Improving the performance of autonomous driving through deep reinforcement learning,"*Sustainability*, vol. 15, no. 18, Art. no. 13799, 2023.

[7] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. Al Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey,"*IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 4909–4926, 2021.

[8] D. Guo, S. He, and S. Ji, "Intersection decision making for autonomous vehicles based on improved PPO algorithm,"*IET Intell. Transp. Syst.*, vol. 18, pp. 2921–2938, 2024.

[9] F. Poszler, M. Geisslinger, J. Betz, and C. Lütge, "Applying ethical theories to the decision-making of self-driving vehicles: A systematic review and integration of the literature,"*Technol. Soc.*, vol. 75, p. 102350, 2023.

[10] G. P. Antonio and M.-D. Cáceres, "Multi-agent deep reinforcement learning to manage connected autonomous vehicles at tomorrow's intersections,"*IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7033–7043, 2022.

[11] H. Karacalı, E. Cebel, and N. Donum, "Acceleration/Deceleration Detection System—Creation of Driving Behavior Profiles for Efficiency,"*The Eurasia Proceedings of Science Technology Engineering and Mathematics*, vol. 26, pp. 519–531, 2023.

[12] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions,"*SN Computer Science*, vol. 2, no. 3, Art. no. 160, 2021.

[13] J. Chen, C. Zhao, S. Jiang, X. Zhang, Z. Li, and Y. Du, "Safe, efficient, and comfortable autonomous driving based on cooperative vehicle infrastructure system,"*Int. J. Environ. Res. Public Health*, vol. 20, no. 1, Art. no. 893, 2023.
[14] J. Guo and I. Harmati, "Lane-changing system based on deep Q-learning with a request–respond mechanism,"*Expert Syst. Appl.*, vol. 235, Art. no. 121242, 2024.
[15] J. Hu et al., "A survey of decision-making and planning methods for self-driving vehicles,"*Front. Neurorobotics*, vol. 19, p. 1451923, 2025.
[16] J. Lin, P. Zhang, C. Li, Y. Zhou, H. Wang, and X. Zou, "APF-DPPO: An automatic driving policy learning method based on the artificial potential field method to optimize the reward function,"*Machines*, vol. 10, no. 7, p. 533, 2022.
[17] J. Markowitz and E. W. Staley, "Clipped-objective policy gradients for pessimistic policy optimization,"*arXiv preprint*, arXiv:2311.05846, 2023.
[18] J. Xie, Y. Qin, Y. Zhang, T. Chen, B. Wang, Q. Zhang, and Y. Xia, "Towards human-like automated vehicles: Review and perspectives on behavioural decision making and intelligent motion planning,"*Transportation Safety and Environment*, vol. 7, no. 1, Art. no. tdae005, 2025.
[19] K. Lv, X. Pei, C. Chen, and J. Xu, "A safe and efficient lane change decision-making strategy of autonomous driving based on deep reinforcement learning,"*Mathematics*, vol. 10, no. 9, p. 1551, 2022.
[20] M. Aizat, N. Qistina, and W. Rahiman, "A comprehensive review of recent advances in automated guided vehicle technologies: Dynamic obstacle avoidance in complex environments toward autonomous capability,"*IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–25, 2023.
[21] M. Bellone, A. Ismailogullari, J. Müür, O. Nissin, R. Sell, and R. M. Soe, "Autonomous driving in the real-world: The weather challenge in the Sohjoa Baltic project," in *Towards Connected and Autonomous Vehicle Highways: Technical, Security and Social Challenges*, Cham, Switzerland: Springer Int. Publ., pp. 229–255, 2021.
[22] M. Everett, Y. F. Chen, and J. P. How, "Collision avoidance in pedestrian-rich environments with deep reinforcement learning,"*IEEE Access*, vol. 9, pp. 10357–10377, 2021.
[23] M. Hua, D. Chen, X. Qi, K. Jiang, Z. E. Liu, Q. Zhou, and H. Xu, "Multi-agent reinforcement learning for connected and automated vehicles control: Recent advancements and future prospects,"*arXiv preprint*, arXiv:2312.11084, 2023.
[24] M. M. Haque, S. Sarker, and M. A. A. Dewan, "Driving maneuver classification from time series data: A rule based machine learning approach,"*Appl. Intell.*, vol. 52, no. 14, pp. 16900–16915, 2022.
[25] M. Reda, A. Onsy, A. Y. Haikal, and A. Ghanbari, "Path planning algorithms in the autonomous driving system: A comprehensive review,"*Robotics and Autonomous Systems*, vol. 174, p. 104630, 2024.
[26] M. Soori, A. Barezoo, and R. Dastres, "Artificial intelligence, machine learning and deep learning in advanced robotics: a review,"*Cognitive Robotics*, vol. 3, pp. 54–70, 2023.
[27] M. Towers, A. Kwiatkowski, J. Terry, J. U. Balis, G. De Cola, T. Deleu, and O. G. Younis, "Gymnasium: A standard interface for reinforcement learning environments,"*arXiv preprint*, arXiv:2407.17032, 2024.
[28] P. S. Chib and P. Singh, "Recent advancements in end-to-end autonomous driving using deep learning: A survey,"*IEEE Trans. Intell. Vehicles*, vol. 9, no. 1, pp. 103–118, 2023.
[29] R. Zhang, J. Hou, F. Walter, S. Gu, J. Guan, F. Röhrbein, and A. Knoll, "Multi-agent reinforcement learning for autonomous driving: A survey,"*arXiv preprint*, arXiv:2408.09675, 2024.
[30] S. A. Useche, R. Mora, F. Alonso, and O. Oviedo-Trespalacios, "Sensation seeking and crashes among young cyclists,"*Accident Analysis and Prevention*, vol. 214, Art. no. 10797, 2025.
[31] S. Malik, M. A. Khan, and H. El-Sayed, "Carla: Car learning to act—an inside out,"*Procedia Comput. Sci.*, vol. 198, pp. 742–749, 2022.

[32] S. Rezwana and N. Lownes, "Interactions and behaviors of pedestrians with autonomous vehicles: A synthesis,"*Future Transportation*, vol. 4, no. 3, pp. 722–745, 2024.
[33] S. V. Bondre, B. Thakre, U. Yadav, and V. D. Bondre, "Deep reinforcement learning algorithms: A comprehensive overview," in *Deep Reinforcement Learning and Its Industrial Use Cases: AI for Real-World Applications*, pp. 51–73, 2024.
[34] T. Zhang, H. Liu, W. Wang, and X. Wang, "Virtual tools for testing autonomous driving: A survey and benchmark of simulators, datasets, and competitions,"*Electronics*, vol. 13, no. 17, Art. no. 3486, 2024.
[35] W. H. Chen, "Perspective view of autonomous control in unknown environment: Dual control for exploitation and exploration vs reinforcement learning,"*Neurocomputing*, vol. 497, pp. 50–63, 2022.
[36] W. Terapaptommakol, D. Phaoharuhansa, P. Koowattanasuchat, and J. Rajruangrabin, "Design of obstacle avoidance for autonomous vehicle using deep Q-network and CARLA simulator,"*World Electr. Veh. J.*, vol. 13, no. 12, Art. no. 239, 2022.
[37] Y. Chen, C. Ji, Y. Cai, T. Yan, and B. Su, "Deep reinforcement learning in autonomous car path planning and control: A survey,"*arXiv preprint arXiv:2404.00340*, 2024.
[38] Y. Du et al., "Comfortable and energy-efficient speed control of autonomous vehicles on rough pavements using deep reinforcement learning,"*Transp. Res. Part C: Emerging Technol.*, vol. 134, p. 103489, 2022.
[39] Z. Bi, X. Guo, J. Wang, S. Qin, and G. Liu, "Deep reinforcement learning for truck-drone delivery problem,"*Drones*, vol. 7, no. 7, Art. no. 445, 2023.
[40] Z. Liu, X. Xu, P. Qiao, and D. Li, "Acceleration for deep reinforcement learning using parallel and distributed computing: A survey,"*ACM Comput. Surv.*, vol. 57, no. 4, pp. 1–35, 2024.
[41] Z. Tian, Z. Lin, D. Zhao, W. Zhao, D. Flynn, S. Ansari, and C. Wei, "Evaluating Scenario-based Decision-making for Interactive Autonomous Driving Using Rational Criteria: A Survey,"*arXiv preprint* arXiv:2501.01886, 2025.
[42] Z. Zhang, E. Demir, R. Mason, and C. Di Cairano-Gilfedder, "Understanding freight drivers' behavior and the impact on vehicles' fuel consumption and CO_2e emissions,"*Oper. Res.*, vol. 23, no. 4, pp. 59, 2023.
[43] Z. Zhu and H. Zhao, "A survey of deep RL and IL for autonomous driving policy learning,"*IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14043–14065, 2021.
[44] Z. Zhu and H. Zhao, "A survey of deep RL and IL for autonomous driving policy learning,"*IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 14043–14065, 2021.