INTERNATIONAL JOURNAL OF AUTOMOTIVE AND MECHANICAL ENGINEERING
ISSN: 2229-8649 e-ISSN: 2180-1606
VOLUME 23, ISSUE 2, 2026, 13691 - 13715
DOI: https://doi.org/10.15282/ijame.23.2.2026.16.1035

REVIEW ARTICLE

A comprehensive review of machine learning approaches for decision-making in autonomous vehicles collision avoidance systems

Siti Nurhafizza Maidin[1], Nurbaiti Wahid[1]*, Sarah Atifah Saruchi[2], Sumiya Tamanna Fujita[2], Alvi Khan Chowdhury[3]

[1]Faculty of Electrical Engineering, Universiti Teknologi MARA, 23000 Dungun, Terengganu, Malaysia
[2]Faculty of Manufacturing and Mechatronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Malaysia
[3]School of Engineering and Advance Engineering Platform, Monash University Malaysia, Selangor, Malaysia

Abstract - Autonomous vehicles (AVs) hold the potential to reduce road accidents statistics through intelligent collision avoidance systems. This paper presents a comprehensive review of machine learning (ML) approaches for decision-making in AV collision avoidance. The comprehensive review is systematically categorized into four categories which are supervised learning (SL), hybrid learning (HL), deep learning (DL), and reinforcement learning (RL). Each category is evaluated in terms of its applicability in decision-making tasks, which include perception for situational awareness task, behavioural-level decision-making task, motion planning task, and multi-agent coordination task. The review highlight that SL methods offer reliable performance in structured scenarios but have constraints under real-world uncertainties. For HL, the methods utilized enable for feature selection and risk prediction improvement but still face challenge of computational overhead. Meanwhile, DL models advance in perception and control enhancement through spatial-temporal reasoning but encounter large-scale data and high processing power demand issue. As for RL, the methods demonstrate advantages of strong adaptability and scalability, supporting multi-agent coordination and decision-making under uncertainty. However, they still face challenges in training efficiency and real-time responsiveness. Based on the findings, the results offer a thorough guide for creating ML-enabled collision avoidance frameworks that are secure, scalable, and reliable for autonomous driving applications. This review also emphasizes existing research challenges such as computational inefficiency, restricted generalization, sensor blockage, and ethical uncertainty in unavoidable collision situations. The review concludes by stating the future research directions toward hybrid architectures, lightweight and edge-deployable models, uncertainty-aware reasoning deployment, and ethically integrated decision-making frameworks. This review provides a structured classification and comparative analysis to guide researchers and practitioners in advancing safer, more robust and socially acceptable autonomous driving systems.

Article History
Received : 15 June 2025
Revised : 28 January 2026
Accepted : 12 February 2026
Published : 30 June 2026

Keywords
Autonomous vehicles
Decision-making tasks
Collision avoidance
Machine learning approaches
Deep learning
Reinforcement learning

1. Introduction

The automotive industry is deeply concerned about traffic safety and road accidents due to the rising number of crashes. Based on [1], traffic accidents are now one of the leading causes of mortality worldwide. The three major factors that contribute to traffic accidents are automobiles, road conditions, and human error [2]. This traffic accident data is from the Malaysian government's open data portal and was analysed for the period 2010 to 2017. According to a study [3], nearly 90% of automotive accidents are caused by human error. According to Wang et al. [4], 94% of accidents are caused passively by other parties, such as pedestrians, cyclists, motorcyclists, and traditional cars, while 6% are directly attributed to AVs. Self-driving cars [5], also known as autonomous vehicles (AVs) [6], have the potential to dramatically transform the experience of travelling and commuting. AVs are designed to eliminate the inherent flaws of human drivers by relying on advanced technologies such as artificial intelligence (AI), machine learning (ML), sensors, and high-definition (HD) mapping. The three main constituents of AVs are perception, decision-making, and motion control systems [7]-[9], as shown in Figure 1. This three-layer hierarchy serves as the operational architecture for AVs, enabling them to drive more effectively. The perception system functions as the eyes and ears for an AV. It consists of sensors such as cameras, LiDAR, RADAR, and ultrasonic sensors that continuously collect data about the vehicle's surroundings. This comprehensive sensory input enables the car to detect obstacles, pedestrians, traffic signs, and other vehicles in real-time. Meanwhile, decision-making is considered the brain of an AV [10]-[11]. It is used to process this perceptual data to reason, predict the behaviour of other agents, and make high-level behavioural choices, such as lane changes, braking, and overtaking. According to Badue et al. [5], the decision-making system is typically divided into multiple subsystems, each responsible for specific tasks such as route planning, path planning, behaviour selection, motion planning, and vehicle control. In the control layer, it acts as the muscles of AVs, converting these choices into accurate steering, acceleration, and braking commands to carry out the intended course.

The significance of decision-making [12] in AVs cannot be overstated. The impact of each actuator can reduce the percentage of collisions or, conversely, make them more catastrophic. This important component must be thoroughly understood and carefully considered. In situations where an accident is unavoidable, the vehicle's decision-making module must determine who or what to prioritise for safety. The decision-making system is one of the most challenging

*CORRESPONDING AUTHOR | Nurbaiti Wahid | ✉ nurbaiti@uitm.edu.my

components in implementing a trustworthy and secure autonomous avoidance manoeuvre when collision risk is present. In particular, decision-making remains largely overlooked in collision avoidance strategies, as evidenced by its inclusion in only 12% of the reviewed research [13]. Since decision-making is a crucial component of autonomous driving systems, it has received significant attention and remains one of the most significant issues in Intelligent Transportation Systems (ITS) research [14]. This part of the system requires significant effort to produce an effective collision avoidance system (CAS). The main goal of CAS is to prevent accidents and ensure the safety of vehicle occupants and other road users. Therefore, it is crucial to implement an effective and secure manoeuvre for successful and advanced CAS. The decision-making system processes the vast amount of data collected by the sensing system. Using complex algorithms and artificial intelligence techniques, it analyses this information to make informed decisions about how to navigate through different scenarios. This includes determining when to accelerate or decelerate, when to change lanes or make turns, and how to respond to unexpected events on the road. Understanding the unknowns and selecting the best driving strategy for AVs is the primary challenge in the autonomous decision-making process. Decision-making in AVs plays a pivotal role in determining their safety and effectiveness on the roads. Decision-making is formulated to generate human-level safe and reasonable driving behaviours, considering surrounding environmental information, the motion of other traffic participants, and state estimates of ego vehicles. Then, the generated driving behaviours are accounted for by the motion control system to achieve efficient operation of the autonomous vehicle [15]-[16]. High-level autonomous driving depends on the decision-making system, which primarily consists of risk assessment, trajectory planning, motion planning, and behaviour decision-making modules. It is considered the brain of an AV [10]-[11].

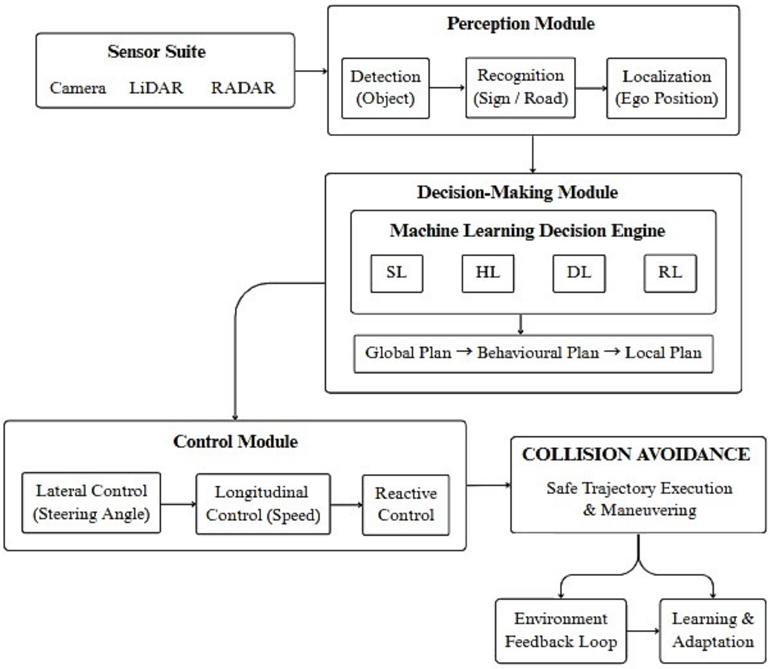

Figure 1. Conceptual framework for autonomous vehicle collision avoidance using machine learning-based decision-making [9]

The current challenge in the real-time decision-making algorithm for CAS is the dynamic environment [17], which includes multiple traffic participants, such as other incoming vehicles, cyclists, and pedestrians. When vehicles operate in mixed-traffic environments, where the actions of surrounding agents create dynamic and inherently unexpected conditions, the decision-making dilemma in autonomous driving worsens. The future state of the environment must be accurately predicted to achieve this goal. To accurately predict the AVs' surrounding environment, sensor utilisation supports perception. For example, camera sensors are mainly used to capture visual information and use computer vision algorithms to recognise barriers, cars, pedestrians, and traffic signs. The LiDAR sensors utilise laser beams to generate high-resolution point clouds, enabling precise distance measurements and accurate environmental modelling. The radar sensors deploy radio waves to determine the location and speed of nearby objects. As a result, these sensing modalities provide the complementary information AVs require for accurate perception and decision-making. The possible actions AVs can take to avoid collisions include lane changes, deceleration, and acceleration. These actions demonstrate their superior safety capabilities compared to those of human drivers. For lane changes, it enables AVs to avoid potential obstacles or hazardous conditions. Meanwhile, deceleration enables AVs to slow down in the event of sudden stops or unexpected obstacles in their path. As for acceleration, it allows AVs to swiftly move out of the danger zone or accelerate to prevent a collision. These actions demonstrate how AVs can proactively prevent accidents while keeping all road users safe. Figure 2 shows an example of a multiple-vehicle collision that can occur, including a regular driving event, a pre-collision event, a 1st collision event, and a multiple-collision event.

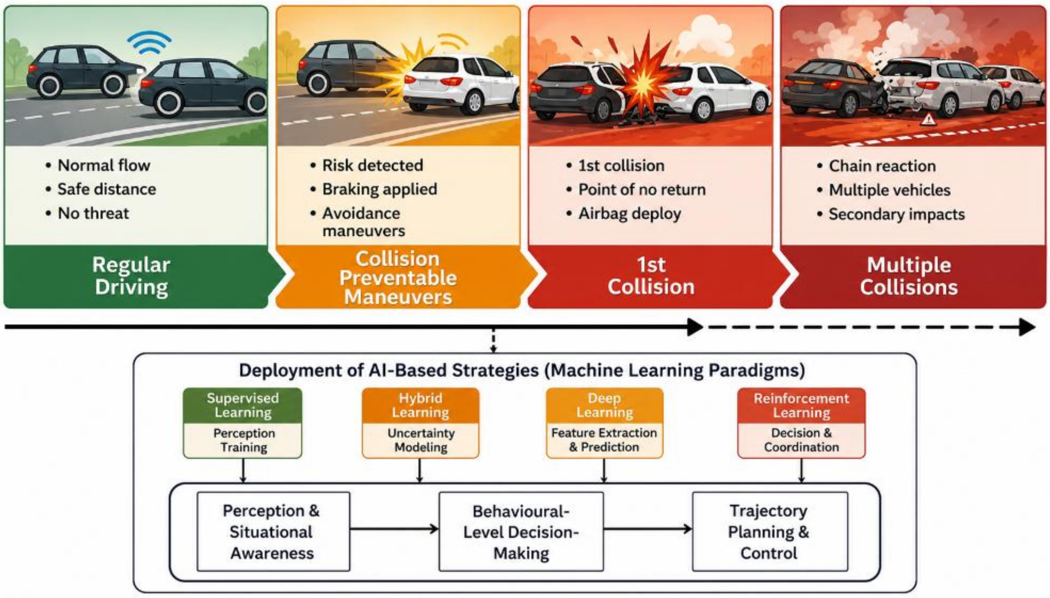

Figure 2. Multiple vehicle collision that can happen, which includes regular driving event, pre-collision event, 1st collision event and multiple collisions event [18]

Machine learning-based classification serves as a crucial bridge, converting raw sensor data into useful information for decision-making in AVs Classification in AVs' decision-making does not just involve labelling data; it also involves classifying driving scenarios, detecting objects, and evaluating dangers. The decision-making category serves as the fundamental input for the behavioural planning and control modules of AVs. The main challenge for AVs is designing approaches that can entirely replace human driver decision-making without compromising safety [19]. Machine learning plays a crucial role in the decision-making process of Advanced Driver Assistance Systems (ADAS) [20]-[21]. In the beginning, it helps the system to interpret real-time data such as nearby obstacles, the vehicle's speed, and distance to other objects. The interpretation is then used to predict potential collisions and apply emergency braking automatically if necessary. It also helps the vehicle adapt to changing traffic and road conditions by adjusting its cruising speed and following distance, ensuring a smoother, more fuel-efficient ride. ML models assist with lane-keeping by analysing the car's position relative to the road and detecting signs, thereby helping prevent unintended lane departures. Furthermore, ADAS also leverages ML to monitor blind spots, alerting drivers to nearby vehicles and reducing the risk of side collisions. Lastly, the system enables smarter driving decisions, either when to slow down, proceed or enhance both safety and efficiency, by recognising traffic lights and road signs [22].

Figure 3 highlights a range of ML approaches for decision-making tasks in an AV's avoidance system. This figure summarises the utilisation of supervised learning (SL), hybrid learning (HL), deep learning (DL) and reinforcement learning (RL) on different decision-making tasks. In [23], supervised learning is defined as the learner producing predictions for every point that is not visible after receiving a collection of labelled samples as training data. In other words, supervised learning is used to train on a labelled dataset and learn the mapping between inputs (features) and known outputs (labels) to make predictions on unseen data. Hybrid learning refers to the integration or combination of different learning paradigms to leverage their respective strengths. In some contexts, it can represent the simple combination of supervised learning [24] and unsupervised learning [25]. Deep Learning has emerged as an essential technology in ML, AI, and data science. This is due to its exceptional ability to extract meaningful patterns and insights from complex datasets. Its capacity for autonomous feature learning and hierarchical representation has positioned DL as a transformative approach across these disciplines [26]. Deep learning uses a series of neural network layers to extract features hierarchically, with each layer recognising and understanding more complex patterns from the initial raw data. This approach employs specialised architectures, including convolutional neural networks (CNNs) for spatial data processing, recurrent neural networks for sequential data analysis, and deep neural networks (DNNs) for general pattern recognition. Among these, convolutional networks are particularly prominent for their effectiveness in processing structured grid-like data such as images [27].

Reinforcement learning is a machine learning technique that maximises the cumulative reward [28]-[29]. The objective of RL is to balance exploration and exploitation. It is mainly because the agent is learning from the environmental information. The foundation of RL is trial-and-error exploration with a delayed reward that aims to learn the best course of action to maximise the long-term rewards for future actions. The fundamental model of RL is the Markov Decision Process (MDP), represented by the tuple $\langle S, A, P, R, \gamma \rangle$. These letters denote the state space, action space, state transition probability (or state transfer matrix), reward function, and discount factor, respectively [30]-[31]. Recently, RL has shown exceptional success in artificial intelligence due to its distinct feedback mechanisms and trial-and-error learning [32]. In general, most AV decision-making modules are rule-based and have advanced significantly in settings such as highways, semi-closed parks, and urban roadways. Nevertheless, rule-based approaches remain inflexible and lack intelligence in real-world applications, and they are computationally demanding. Therefore, the adoption of ML in decision-making tasks advances AV technology by enabling AVs to better understand the environments in which they operate.

The primary focus of this review is the AV's decision-making layer (with an emphasis on obstacle avoidance), which connects perception and control. To ensure safe navigation, the decision-making module regulates reasoning, prediction, and behavioural choice, while perception facilitates environmental understanding and control guarantees vehicle actuation. Nevertheless, this review will also cover important perception techniques such as object detection and depth estimation, which are directly or inherently related to the subsequent decision-making process. This is because the high-quality decisions depend on an accurate understanding of the environment. This review also includes ML approaches on different decision-making tasks for collision avoidance in AVs. This review will summarise how the proposed approaches of SL, HL, DL and RL function in perception, situational awareness, behavioural decision-making, motion planning, and multi-agent decision-making. This comparative analysis includes the proposed method, key contributions to decision-making, key metrics of the approaches, reported performance, test environments, and the approaches' limitations. This paper is organised as follows: ML approaches for decision-making tasks in collision avoidance in Section 2; a comparative analysis of all approaches in Section 3; and conclusions in Section 4.

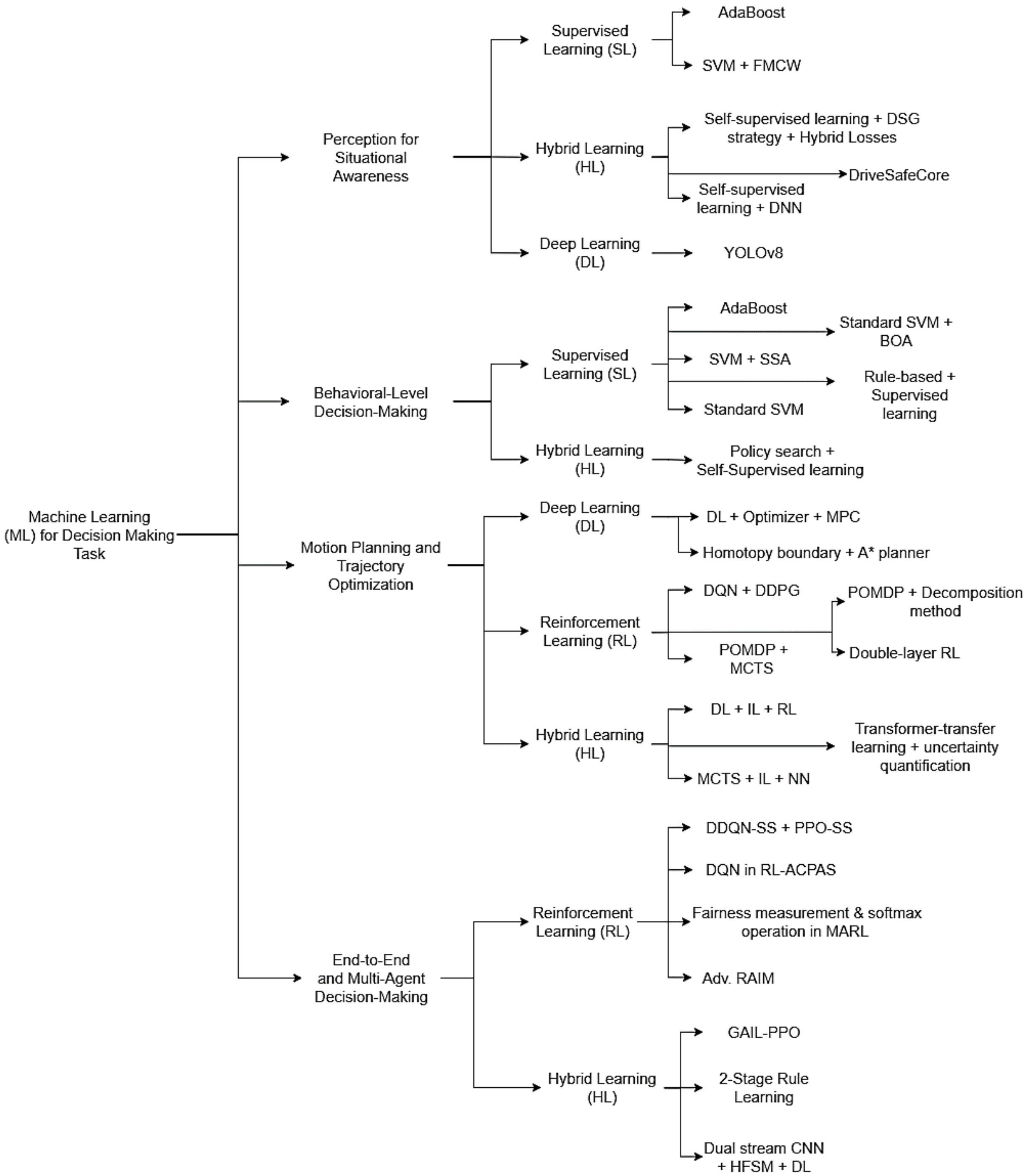

Figure 3. Artificial intelligence and machine learning categorised based on the autonomous vehicle system tasks

2. Machine Learning for Decision-Making Tasks in Collision Avoidance

Here, the different ways ML paradigms address specific decision-making tasks, from interpreting sensor data to planning high-level manoeuvres, are analysed.

2.1 Perception for Situational Awareness

According to Long et al. [33], this approach effectively filters out a significant amount of visual input that does not contain vehicles, thereby aiding real-time decision-making. Negative samples, or regions of the image devoid of vehicles, were promptly identified using the easy, fast classifiers in the initial stages. Rapid decision-making is possible by thorough examination of existing vehicle regions. Kalman filtering makes the tracking decision-making process more precise and dependable by combining historical position data with present observations to forecast future vehicle movements. In Chaudhary et al. [34], the SVM method and Frequency-Modulated Continuous-Wave (FMCW) photonic radar [35]-[36] are integrated to efficiently classify various targets in challenging situations and to enhance the system's ability to detect potential collisions. To improve classification performance, two SVM models were used for different purposes. One model is dedicated to classifying stationary targets, while another is developed specifically for classifying moving targets. This method contributes to the classification of a variety of targets in decision-making in AVs and complex scenarios. As stated in Zhao et al. [37], integrating the hybrid loss function with the Depth Self-supervised Generative (DSG) strategy offers a better approach to depth estimation for AVs and provides a better outlook for the decision-making module. The proposed implementation of a hybrid loss function, which includes photometric projection loss, automatic masking loss and smoothness loss, helps to minimise the Structural Similarity Index [38]-[39] and L1 norm error [21] to predict depth accurately, enhance depth prediction visual quality in dynamic scene changes and improve depth prediction visual quality. The strategy of DSG is to adapt Gaussian noise, speckle noise and dynamic fuzzy to enhance robustness for noise-augmented training data. Consequently, the proposed methods achieve 8.2% higher accuracy on noisy data while maintaining performance on clean data. In the paper by Stojcheski et al. [40], integration of self-supervised learning and a deep neural network approach in occupancy grid maps (OGM) [41] enables improved data mapping in autonomous driving. The proposed approaches help design the map based on incomplete information obtained from the autonomous vehicle's perception. The utilisation of temporal aggregation in a self-supervised approach enables the generation of real-world training data in distant and occluded areas. The adaptation of deep neural network architectures, such as SegNet [42], U-Net [43]-[44] and Full-Resolution Residual Network (FRRN) [45], also contributes to OGM completion. Correspondingly, the layout of the distant or occluded areas able to be mapped can substantially enhance degraded OGM observations in regions affected by distance and occlusion.

In Sharma et al. [46], YOLOv8 [47]-[48] is utilised via transfer learning, and the Canadian Vehicle Dataset (CVD) is introduced, gathered from real-world driving in Quebec under various weather conditions. An autonomous vehicle's ability to accurately comprehend complex environments is directly improved by the stated gains, which include a 29.1% increase in precision and a 33% increase in recall, resulting in a mean Average Precision (mAP) of 0.475. For the following decision-making layer to enable safer, better-informed navigation, collision avoidance, and path planning in difficult real-world settings, this dependable object detection is essential. The study in Beg et al. [49] describes a real-world application of a YOLOv8-based perception system for driving assistance, verified on Malaysian roads. Reliable object recognition is a crucial first step in AVs' decision-making process because it provides vehicles with the environmental awareness needed to make informed navigation and collision-avoidance decisions. However, a limited number of object classes and a tiny, regionally localised dataset restrict the study's breadth, potentially undermining the decision-making system's resilience in more complex or varied driving settings. The study effectively illustrates how deep learning can improve the perceptual underpinnings of autonomous driving systems despite these drawbacks. Based on the paper by Duan et al. [50], the method used effectively improves the precise identification and categorisation of different road objects, which is necessary for safe driving. This method enables unsupervised clustering to identify and categorise features across different datasets into meaningful groupings that serve as pseudo-labels for training. Then, using the cascade AdaBoost structure, the clustered data was refined, and a combination of weak classifiers was created via AdaBoost. Thus, AVs will achieve higher prediction accuracy and make decisions more efficiently.

Supervised learning and unsupervised learning are integrated in Chen et al. [51], allowing object detection robustness in AVs. Transfer learning based on the Momentum Contrast (MoCo) [52], semi-supervised co-training using MoCov1 and MoCov2 [53] and Bounding Box Augmentation (BBAug) implementation contribute to enhancing the capability of current object detection systems to withstand adversarial threats. The adversarial attacks include the Dpatch [54] and Contextual Patch [55] white-box methods, as well as Image Corruptions black-box attacks. The BBAug makes the detector more resilient in challenging environments and against adversaries by serving as a robustness enhancer added during the retraining step. As a result, the MoCo-based transfer learning serves as a robust feature extraction backbone, while the semi-supervised co-training makes it easier to transfer knowledge from labelled to unlabelled data using high-confidence pseudo-labels produced during inference. The DriveSafeCore is proposed in Prabha et al. [56], which includes the integration of multisensory fusion, spatio-temporal scene understanding, modelling of trajectory prediction and DRL. The proposed method enables in-depth understanding of the scene using sensor fusion data, leading to precise trajectory prediction. The utilisation of DRL in the proposed method enables the ego vehicle to make optimal decisions, as the DQN will develop a driving strategy that generalises to unfamiliar environments and unpredictable traffic scenarios. Aside from that, the redundant data obtained from sensor fusion will be combined, and accidental fluctuations will be removed through Kalman filtering. This will provide high-quality data to the decision-making module to design decisions that suit the encountered environment. Using the proposed method, the ego vehicle can avoid collisions with surrounding vehicles in the simulated environment in real time, and an efficient decision-making module can be designed. Improvements in perception, from robust sensor fusion and effective feature filtering to complex self-supervised depth estimation and high-accuracy real-time object detection, provide the essential environmental understanding layer that enables subsequent decisions in AVs Table 1 provides a simple analysis of the methods and key contributions to decision-making in the situational awareness perception task. Although accurate perception is crucial for developing a reliable representation of the driving surroundings, advanced cognitive processing is required to convert this awareness into safe and rational vehicle manoeuvres. The decision-making layer is responsible for interpreting the detected objects, predicted trajectories, and evaluated hazards to provide tactical driving actions, such as lane changes, overtaking manoeuvres, and speed control, to the AVs This section discusses behavioural-level decision-making techniques from the perspectives of supervised learning, hybrid learning, and deep learning. The designed ML models are required to convert perception outputs into sequential, discrete driving policies that concurrently prioritise human-like interaction, operational efficiency, and safety in dynamic traffic environments.

Table 1. Methods and key contributions to decision-making of the perception task

Methods	Key Contributions to Decision-Making
AdaBoost + Kalman Filter [33]	AdaBoost enables unnecessary visual data to be filtered out, which allows quick recognition of vehicles. Kalman filter enhances tracking precision and reduces the delay between receiving sensor data and executing actuator actions
SVM + FMCW [34]	High-accuracy target detection reduces false positives and negatives, leading to a decrease in the possibility of missed detections or incorrect classifications. The FMCW allow for rapid data processing in real time.
Self-supervised learning + DSG strategy + Hybrid Losses [37]	The proposed algorithm provides greater speed and enhanced precision, ensuring perception data is processed with minimal latency, leading to accurate detection of traffic participants and reducing uncertainties
Self-supervised learning + Deep neural network [40]	Accuracy and inference speed in the perception algorithm can improve understanding in occluded or distant regions, aiding safer and more confident decision-making in autonomous driving
YOLOv8 [46]	Offers a more accurate and dependable perceptual foundation for object detection, especially when trained on a varied and weather-specific dataset
YOLOv8 [49]	High detection accuracy and real-time performance offer a dependable perceptual foundation to enable prompt and secure decision-making in autonomous driving systems
Deep Cascade AdaBoost + Unsupervised Clustering [50]	Detection time reduced by 30% in dynamic environments, and the accuracy of object identification improved for safer driving
Transfer learning + Co-training + Data Augmentation [51]	Enhances object feature learning using unlabelled data, improves generalisation by adding high-confidence unlabelled data and increases robustness to noise distortions
Data Fusion + Temporal awareness + RL in DriveSafeCore [56]	Enable optimal decision provided for collision avoidance event in real-time with the help of the multimodal sensor input fusion, temporal awareness, trajectory prediction and DRL

2.2 Behavioural-Level Decision-Making

In Singh et al. [57], the boosting [58]-[59] method was implemented under conditions where decision-making time was limited to 1, 2, or 3 seconds. The 1-second condition required immediate decision-making in response to unfolding events, whereas the 2-second condition encouraged the assimilation of more information and the re-evaluation of initial judgements. As for the 3-second condition, it will provide optimal predictions for AVs The Adaboost [60] demonstrates superior performance in ethical decision-making for AVs, as it excels at achieving high true positive rates and controlling false positives. Adaboost can also perform well under stringent time constraints, ensuring that AVs can make ethical choices that prioritise human safety. The optimisation method, the Sparrow Search Algorithm (SSA), [61], is utilised to fine-tune the SVM's parameters. Integrating the mentioned method with game theory, which enables it to account for the strategies and potential reactions of nearby vehicles, leads to better performance in vehicle lane-changing decision-making. For an amalgam of SVM and Bayesian Optimisation Algorithm (BOA) in Liu et al. [62], autonomous vehicles perform lane change in terms of decision-making based on benefits, safety, and tolerance. These three aspects are considered when performing a lane change to ensure smooth execution involving the ego vehicle, the preceding vehicle in the target lane, the rear vehicle in the target lane, and the preceding vehicle in the original lane. Once the benefits and safety value are high enough, the tolerance value is applied to execute a safe lane change. Aside from that, BOA also provides efficiency improvements in C and kernel parameters for lane change decision-making in autonomous vehicles. The SVM method [63] is implemented in the autonomous vehicle [64] to support the lane-keeping assist system. This method is used to determine the steering angle of an autonomous vehicle based on sensor input and to keep the vehicle always centred in the lane, under both static and dynamic conditions. This method determines which direction (left or right) the vehicle should take. The positive classes (right) and negative classes (left) are split across the hyperplane to determine the steering angle applied to the autonomous vehicle. Based on Wang et al. [65], augmenting a dynamic decision model with artificial neural networks (ANNs) [66]-[67] and control algorithms for collision-avoidance systems on a vehicle helps avoid collisions with obstacles and with jamming vehicles in a two-lane expressway scenario. The dynamic decision model generates preliminary braking and steering commands using real-time traffic, vehicle dynamics, and environmental data. The model also provides the vehicle with two collision-avoidance modes: braking collision avoidance and coordinated collision avoidance (braking and steering). The outputs of the dynamic decision model, which are maximum braking deceleration and maximum lateral acceleration, are then refined using the ANN controller. Therefore, the model determines which type of collision avoidance manoeuvre the vehicle should execute, based on real-time parameters that lead to effective coordinated collision avoidance in AVs

Based on Laneve et al. [68], the integration of weighted maximum-likelihood policy search and self-supervised learning contributes to lane-change manoeuvres in dynamic, online adaptation among the ego vehicle, the lateral vehicle, and the front vehicle. The upper-level policy determines the optimal timing for the vehicle to perform the lane change with the help of the Model Predictive Control (MPC) system, and self-supervised learning keeps pace with any changes, intentional or unintentional, in the environment. Monte Carlo Expectation Maximisation (MC-EM) [69]-[70] adaptation in MPC allows for critical time-switching (θ) variable optimisation and supports adjustment on how much the AV prioritises staying in the current lane (ego cost) versus moving to the target lane (target cost) in its trajectory planning. As a result, the ego vehicle can overtake the front vehicle before or after the lateral vehicle passes, with collision-free travel between them. In Wang et al. [71], the Long Short-Term Memory (LSTM) [72]-[73] method and the Adaptive Moment Estimation (Adam) [74] optimiser were utilised. To analyse the lane-change decision-making intention in the ego vehicle based on the sequence of input data related to the ego vehicle and the surrounding environments. As LSTM is used in hierarchical decision-making for lane-change intention, it requires evaluating patterns and producing suitable lane-change intentions for the ego vehicle. Then, the Adam optimiser effectively refines the model to capture intricate patterns in the data, ensuring it converges more rapidly and precisely. A stage decision-making framework is used in [32]. The framework consists of a route-planning level and a micro-level of car-following (CF) and lane-change (LC) behaviour. This framework utilises Deep Q-Network (DQN) [75]-[76] and Deep Deterministic Policy Gradient (DDPG) [77]-[78] algorithms. DQN is used to discretise and learn lane-changing behaviour, while DDPG is used to learn continuous car-following behaviour. The implementation of the reward function in these algorithms contributes to strategies that reach the destination faster and make more stable decisions than humans.

Table 2. Methods and key contributions to the decision-making of the behavioural decision-making

Methods	Key Contributions to Decision-Making
AdaBoost [57]	AdaBoost trained with experimental data to mimic human-like decision-making in cut-in scenarios and prioritises safety while adapting to varying traffic conditions
SVM + SSA [61]	Integrates SSA with SVM for parameter optimisation and utilises the game theory in SVM to anticipate vehicle interactions, leading to lane-change decisions enhancement in AVs
SVM + BOA [62]	Integrate SVM with BOA for parameter tuning to improve lane-change efficiency, balancing safety and traffic benefits to suit the habits of driver
Standard SVM [64]	Utilise SVM classification to determine appropriate steering adjustments for lane centring based on the input of road images, reference speed and a reference steering angle
Dynamic decision model + ANN + controller [65]	Establishes a robust, real-time classification foundation by learns the braking and steering scale factors to refine control inputs dynamically from traffic or vehicle data
Policy search + Self-Supervised learning [68]	Integrate MPC that improved by adaptive learning for dynamic surroundings with an upper-level policy to eliminate the conservatism of decoupled decision and planning, which responds flexibly to evolving situations and unplanned obstacles
LSTM + Adam Optimizer [71]	LSTM captures patterns in sequential data to predict or mimic the subjective lane-changing intentions of human drivers, and the Adam optimizer refines the model for precise decision-making
Linear algebra-based [79]	The decision is made using matrix operations, and decision thresholds help define vehicle actions
DQN + DDPG [32]	The proposed method allows for continuous lane-change and car-following behaviour with the composite reward function that contributes to comfort, safety and efficiency of AVs
GMA-TD3 [82]	Provide an effective sequential avoidance decision in the multi-ship encounter scenarios

The linear algebra technique is implemented in Darapaneni et al. [79] for decision-making in AVs This technique enables decision-making using matrix operations and a decision threshold. The binary classification of the output images and their conversion from the surroundings contribute to the model's output. Setting the decision threshold to 0.5-1.0 (slow down), below 0.5 (stop the vehicle), and 1.0 (accelerate) helps AVs clearly determine which action to execute and improves the decision-making process. For Shi et al. [80], integration of Fuzzy C-means, XGBoost [81] and recursive feature elimination (RFE) [55], and Bayesian optimisation results in excellent performance, which deals with the behaviour-risk decision-making in AVs As the Fuzzy C-means is used to determine the relative risk levels of various driving vehicles across a wide range of vehicle streams, XGBoost and RFE are used to rank and filter a set of relatively essential features, and to permute the features to identify the best feature combination. Then, the feature optimisation performed by Bayesian optimisation yields more promising values by leveraging prior search results. This combination method provides behaviour-based risk prediction with satisfactory results, with an overall accuracy of 91.7%. Based on [82], the utilised Gated Recurrent Unit-enhanced Multi-head Attention (GMA) helps ownership (OS) identify and classify the collision risk of the surrounding target ship (TS) in real time. The GMA will prioritise obstacles based on their highest and lowest priorities. Once the classification of collision risk for the surrounding TS is completed, the GMA is integrated with Twin Delayed Deep Deterministic Policy Gradient (TD3) to provide a sequential decision to avoid collisions with obstacles. The possible actions executed are ship change course or adjust speed, depending on the situations encountered. The deployment of the dual-level reward mechanism in the Collision Avoidance Decision-Making framework assists the OS in constructing a better strategy to avoid collisions with surrounding OS. The observed behavioural-level strategies show a progression from rule-based classifiers and optimised SVM to tactical actions. Tactical actions, such as lane changes, will help improve deep- and reinforcement-learning frameworks. These improved systems enable flexibility and step-by-step decision-making in dynamic environments with multiple agents. As a result, they create an essential connection between situational awareness and the implementation of motion plans. Table 2 provides a simple analysis of the methods and key contributions to decision-making in the behavioural-level decision-making task. After determining a high-level behavioural decision, such as changing lanes or overtaking, the autonomous system needs to generate a physically feasible, collision-free trajectory that complies with vehicle dynamics and traffic regulations. This conversion from discrete intent to continuous vehicle motion defines the motion-planning and trajectory-optimisation stage, where the emphasis shifts from decision selection to precise execution. Here, the emphasis switches from what has to be done to how to do it safely and smoothly. To navigate complex spatial-temporal settings, the approaches that translate behavioural commands into comprehensive path and velocity profiles are reviewed in the next section.

2.3 Motion Planning and Trajectory Optimisation

A stage decision-making framework is used in [32]. The framework consists of a route-planning level and a micro-level of car-following (CF) and lane-change (LC) behaviour. This framework utilises Deep Q-Network [75]-[76] and Deep Deterministic Policy Gradient (DDPG) [77]-[78] algorithms. DQN is used to discretise and learn lane-changing behaviour, while DDPG is used to learn continuous car-following behaviour. The implementation of the reward function in these algorithms contributes to strategies that reach the destination faster and make more stable decisions than humans. The integration of deep learning, optimisation methods, and model predictive control (MPC) [83] contributes to the realisation of AV decision-making in unknown environments [84]. AV executed manoeuvres in two different scenarios, lane change and obstacle-rich, based on the collision probability output of deep learning, as the raw input is from LiDAR measurements and MPC. The optimisation method chips in by simplifying the objective function and reducing the computational load. Integrating these methods also allows the dynamic collision avoidance problem to be converted into a static one, yielding easier computation and faster solutions in real-time dynamic environments. In Bouton et al. [85], augmenting the decomposition method with the POMDP technique helps utilise the best avoidance technique for a particular user in the occlusion sensor area, which involves a pedestrian crosswalk and intersection scenario. This decomposition strategy involves calculating the optimal state-action utility for each user in the environment, separately from the others, and then combining these utilities to approximate the global problem of avoiding multiple users. The optimal policies that involve a decomposition strategy are Successive Approximations of the Reachable Space (SARSOP) [86]-[87] and the Quantum Markov Decision Process (QMDP) [88]-[89]. Safety and efficiency are used as metrics to evaluate the performance of the optimal policy in those scenarios. As a result, SARSOP and QMDP outperform the random and baseline methods in terms of the average number of collisions and average time to reach the goal position. Based on Gonzalez et al. [90], a partially observable Markov decision process (POMDP) [91] planner implementation, utilising Monte Carlo Tree Search [92] to select the best action for the ego vehicle, improves decision-making during lane changes and longitudinal control. Since the actual state is unknown and the agent receives noisy or incomplete observations of it, it must choose its course of action based on its past performance. The value function for the present belief was approximated during the planning stage using Monte Carlo tree simulations. Therefore, this method enables the ego vehicle to make decisions based on the intentions and ambiguous situations of other cars, replicating how human drivers may predict those cars' intentions. According to Syamil et al. [93], the homotopy boundary method is utilised to classify the area, in the presence of surrounding vehicles, into upper and lower boundary constraints. Using an A* planner in training as the expert planner enables the generation of collision-free longitudinal trajectories for the ego vehicle. The classification boundaries enable smooth trajectory prediction of the ego vehicle. In addition, it creates a safe gap between ego vehicles and surrounding vehicles based on state information and predictions of surrounding vehicles' intentions. The expert demonstrations in the system are used to learn to identify which vehicles fall within the critical upper and lower longitudinal bounds, so that the ego vehicle's safe operation can be determined. Classifying the vehicles ensures that the decision-making module in the ego vehicle is more effective and performs well.

For Arumugam et al. [94], the utilisation of DL, imitation learning (IL), and RL for the decision-making module in AVs enables the generation and use of human-like perception in unseen environments. The structured, filtered processing by the DL generates a rich spatial and semantic understanding of the environment, thereby increasing decision-making efficiency. The IL is used to clone the behaviour of the policy trained on expert demonstrations using supervised learning. The RL is utilised to support higher-level strategic decisions by focusing on desired outcomes and the rationale for multi-step actions. The outputs of each learning-based approach are then ensemble, and the final action is decided using the voting mechanism. This provides an adaptive element in making the optimal decision to execute in an unseen environment. It shows that integrating learning-based approaches into the decision-making module is pivotal, as it leads to better actions during collision avoidance events.

The transformer-transfer learning approach in Liang et al. [95] contributes to predicting the human-driven vehicle (HDV) behaviour and trajectories for AVs. The transformer-transfer learning approach is considered teacher-student learning, as the transformer (teacher) will learn about the HDV lane change maneuverer, then the learning is transferred to the student to enable the trajectory predictions in AVs. The interaction-aware learning between AVs and HDVs alone is insufficient to address trajectory prediction, so uncertainty quantification using the error ellipse method is integrated with the transformer-transfer learning approach. The integration of the proposed method enhances the decision-making and motion-planning module by providing a better strategy for mixed-traffic scenarios and by avoiding collisions when HDVs are present in the environment. The proposed method not only helps AVs learn the behaviour of HDVs and predict and plan HDV trajectories, but also helps them develop an optimal strategy in the presence of uncertainty in trajectory prediction. It shows that integrating the uncertainty quantification method into the motion planning module demonstrates its effectiveness in enhancing AV safety and preventing potential collisions. The proposed method in Chekroun et al. [96] integrates MCTS with a neural network (NN) to improve decision-making and collision avoidance. The realistic world is modelled using the IL, allowing for a simplified representation that includes the 'known features' and 'learned features'. Meanwhile, the NN enable the prediction of the ego and other agents' trajectories present in the environment. Utilising the NN within the environment helps make optimal decisions during MCTS motion planning. MCTS contributes to exploitation and exploration to select the best steering angle and acceleration based on the highest Q-value obtained and nodes with low visitation. By better understanding the environment through a partially learned model and an NN, MCTS can explore and exploit the best nodes to execute. It shows that exploitation and exploration in MCTS, aided by the IL model, are important for finding the best choice among the available options. By considering this method, not only can the trajectories of the ego and agent be predicted, but it also helps plan motion in the best way, enhancing the effectiveness of decision-making and collision avoidance.

Table 3. Methods and key contributions to the decision-making of the motion planning

Methods	Key Contributions to Decision-Making
DQN + DDPG [32]	The proposed method allows for continuous lane-change and car-following behaviour with the composite reward function that contributes to comfort, safety and efficiency of AVs
DL + Optimizer + MPC [84]	Converts dynamic collision avoidance into a static problem for faster and real-time computation by the MPC solver
POMDP + decomposition [85]	Reduces computational complexity and enables scalable and efficient decision-making in dynamic, multi-user environments with the presence of occlusions
POMDP + MCTS [90]	Handles uncertainty in observations by using belief state planning to make decisions under ambiguity, and improves human-like behaviour for lane changes and longitudinal control
Homotopy boundary + A* planner [93]	The classification of upper and lower boundaries in the decision-making module enables the smooth trajectories prediction is generated
DL + IL + RL [94]	Generate the human-like options in unexpected environments and provide an adaptability element in the decision-making module as each learning-based provide their own output to find the optimal decisions
Transformer-transfer learning + uncertainty quantification [95]	Enhancement in the accurate trajectory prediction and collision avoidance, as well as safety features based on the integration of the proposed method for interaction-aware between the AVs and HDV
MCTS + IL + NN [96]	Long-term understanding of the modelled environment helps explicit exploitation and exploration by MCTS for decision-making and collision avoidance enhancement
Double-layer RL [97]	Enhancement in decision based on the feedback provided by the lower layer and employment of trajectory-level reward function evaluated the decision provided, leading to better decision execution

The proposed double-layer RL framework consists of an upper layer and a lower layer, utilised to optimise the decision-making and motion-planning module in Liao et al. [97]. In the deployed closed-loop mechanism, the output from the decision-making module (upper layer) guides the motion planning module (lower layer) to execute an optimal trajectory, including steering and acceleration values. The decision provided by the upper layer is evaluated using a trajectory-level reward function to ensure it is suitable for execution in the lower layer. Then, the trajectory output in the lower layer is fed back to the upper layer for parameter optimisation. Hence, the upper layer can provide better guidance based on the lower layer's feedback. It shows that these 2 modules need to work together, not separately. This double-layer RL framework demonstrates how the combination of these modules contributes to effective decision-making and motion planning, since both modules rely on each other's outputs. When working together, better results can be achieved, and AV safety can be enhanced, particularly during collision avoidance. Together, these motion planning methods illustrate the importance of advanced AI techniques in transforming tactical decisions into safe, executable, and human-like paths. Techniques such as hybrid DRL-MPC controllers and probabilistic POMDP-MCTS planners are crucial, particularly in situations with occlusions and dense traffic interactions. Table 3 provides a simple analysis of the methods and key contributions to decision-making in the motion planning task. Motion planning acts as a crucial layer of translation between executable control and strategic aim. The reviewed techniques demonstrate how sophisticated AI can produce trajectories that are safe, fluid, and dynamically feasible even in the face of uncertainty and obstacles. Real-world driving requires awareness of and cooperation with other intelligent agents, whereas motion planning focuses on creating optimal routes for a single vehicle. In the next section, the sophisticated frameworks which are used to improve contextual adaptation and system-wide coordination beyond conventional single-agent techniques are discussed.

2.4 End-to-End and Multi-Agent Decision-Making

In Gao et al. [98], the implementation of semantic segmentation [99] on the Double Deep Q-Network (DDQN) [100] and Proximal Policy Optimisation (PPO) [101] algorithms helps AVs achieve clearer vision and better driving decisions. The semantic segmentation model designation [99], [102] helps strike a balance between accuracy and efficiency in autonomous driving scenarios. These methods are called Double Deep Q-Network-Semantic Segmentation (DDQN-SS) and Proximal Policy Optimization-Semantic Segmentation (PPO-SS). Semantic segmentation distinguishes and labels scene features, including roadways, traffic signs, pedestrians, and obstacles, by classifying each pixel in a picture. This provided the decision-making algorithm with crucial semantic knowledge to recognise potentially dangerous locations and regions that may be driven through. In Chen et al. [103], the Multi-Agent Reinforcement Learning (MARL) [104]-[105] method was implemented for twin-vehicle cooperative driving in a dynamic and complex highway environment. The deployment of a hybrid framework involves an upper decision-making module that employs a value-based MARL algorithm and a lower control module that optimises speed and orientation. It helps to promote flexible decision-making, feasibility, and safety through lane-keeping and lane-changing models. Besides, each agent can automatically learn a cooperative driving strategy without engaging in unfair behaviour. This is due to the application of a fairness measure that helps balance self-interest and mutual benefit. Moreover, the contributions from the softmax operation help it avoid overestimation when computing Q-values. Using this integrated strategy, the twin cars can balance individual and team goals in dynamic roadway situations and make wise, collaborative judgments.

The Reinforcement Learning–Based Autonomous Collision Prediction and Avoidance System (RL-ACPAS), as implemented in Jagan et al. [106], makes a significant contribution to the field of AV decision-making. The system utilised a DQN to learn collision-avoidance policies in real time. The system uses a learned value function to convert high-dimensional sensory inputs into navigational actions. This enables the crucial problem of autonomous driving to be addressed and yields safe, effective decisions in dynamic traffic settings. Based on Guo et al. [107], hybrid reinforcement learning enables collision avoidance in AVs The proposed method integrates the Soft Actor-Critic (SAC) with feature extraction from a deep neural network to achieve better obstacle detection and appropriate motion to avoid collisions. The method enables effective handling of complex sensory inputs from RADAR and LiDAR, leading to effective decision-making and actuator control during collision avoidance. The proposed method results in a high collision avoidance rate (CAR) and a lower average reaction time (ART) during collision avoidance. According to Ellouze et al. [108], the deployment of Multi-Agent RL (MARL) and the integration of Explainable AI (XAI) have improved the decision-making framework in AVs The 3 specialised agents are utilised in the MARL to break down the large task into critical sub-tasks. The result of the critical sub-task will affect each agent's decision, enabling modular and cooperative learning among them. For XAI implementation, it helps increase human understanding and regulatory compliance by giving each decision a reason for its occurrence. This will increase the transparency of each decision taken and improve the effectiveness of the decision-making module. The Shapley additive explanation (SHAP) score of 0.77 indicates that more collisions can be avoided and that decision traceability for AVs increases.

The incorporation of the attention mechanism and inter-frame redundancy reduction strategy [109] into Generate Adversarial Imitation Learning (GAIL)–PPO reduces long-term redundant input information obtained through the environmental perception module. This method is used to filter out redundant and unnecessary information that would reduce the quality of the collected images. In decision-making, the quality and quantity of images are pivotal because the decision is based on this information; increasing redundancy can improve computational efficiency but also reduce the quality of the input. By focusing only on high-quality input and removing redundant information, the decision-making capability to make efficient decisions is improved, enhancing safety features, particularly in obstacle-avoidance events. The implementation of the reward function design using artificial potential field theory guides AVs to comply with traffic rules, as the reward function includes lane-constraint, obstacle, driving-smoothness, and passing-efficiency potential fields. This shows that the reward function is not only focused on efficiency but also prioritises safety. Deploying the IL model to train the information based on expert prior knowledge for RL decision-making will improve the model's learning efficiency. The performance of end-to-end RL for autonomous driving decision-making can be improved using the proposed method. Wu et al. [110] stated that a 2-stage collective rule-learning framework is deployed in the high-level decision-making module of connected autonomous vehicles (CAVs). In the first stage, the knowledge aggregation will create unified rule-based learning using the Multi-Access Edge Computing (MEC) network. The unified rules are used to accelerate learning by collecting and analysing collision data from multiple CAVs In the second stage, intelligence aggregation assigns weights to the unified rules from the first stage to prevent policy conflicts. The weight assignment in the unified rules will filter out insignificant rules, making the decision model more compact. It shows that the rules with the highest weight are considered the most optimal rules for the decision-making module. Based on the proposed method, knowledge and intelligence aggregation in CAVs is pivotal for providing adaptability, enabling CAVs to adapt to dynamic environments, and enhancing performance in collision avoidance events. This enables collision reduction without requiring expert knowledge. Hence, with the proposed method, the limitations of collision data can be overcome, enabling continuous learning in dynamic driving environments.

The autonomous intersection management (AIM) consisted only of the conflict and priority module. However, the state/conflict encoder is proposed as a module, and the motion planner module, which consists of a fully connected network, is included in the Reinforced AIM (RAIM) system, resulting in advanced RAIM (adv. RAIM) [111]. The proposed state/conflict module enables all vehicle state features to be fed into the LSTM, and the conflict's encoded value is learned during training. The motion planner module determines the adjusted speed for each CAV in the upcoming timestep by considering its characteristics and the state/conflict encoder outputs, ensuring collision avoidance and optimal traffic flow. In addition, the adv. The RAIM approach is based on end-to-end training and uses curriculum learning through self-play. This self-play is used to improve the solution's performance once it is maxed out by adding more data to the training. By implementing the proposed method, CAVs can cross the intersection safely, without collisions, and in a cooperative manner, thereby providing an effective decision-making module for CAVs Towards a fully autonomous driving mode for AVs, decision-making plays a pivotal role in ensuring the safety of vehicles and the user in the best place. To achieve this goal, the integration of a dual-stream convolutional neural network (CNN), a Hierarchical Finite State Machine (HFSM), and a DL module is proposed as the method [112]. These methods enable the encountered scenarios to be classified as simple or complex; HFSM will handle simple scenarios, while DL is used for complex ones. The processed input from the 360-degree front- and rear-view cameras in a dual-stream CNN will assist with situational awareness, as the cross-section between these cameras will indicate whether the situation is simple or complex. The decision-making module, using the proposed method, focuses on determining the driving situation that allows strategy switching during driving. It shows that the adaptability element was present when the proposed method was deployed, as switching occurred when the AVs encountered simple or urban scenarios. Besides, it also provides a generalisation element that can be used in different situations. These advanced frameworks represent the front of autonomous decision-making. These frameworks highlight how end-to-end perception-action learning and multi-agent collaboration are progressing beyond individual vehicle management. This results in scalable, cooperative, and contextually aware driving systems in shared environments. Table 4 provides a simple analysis of the methods and key contributions to decision-making in the multi-agent decision-making task. Integrated, cutting-edge, and cooperative autonomous driving comprises end-to-end and multi-agent frameworks. These methods enable direct mappings from perception to action and encourage socially conscious interactions in shared transportation environments. This is done by extending beyond isolated vehicle control. In the next section, a comparative review evaluates the strengths, limitations, and interdependencies of ML techniques across the four major layers of autonomous vehicle decision-making tasks, from perception to multi-agent coordination.

Table 4. Methods and key contributions to the decision-making of the multi-agent decision-making

Method	Key Contributions to Decision-Making
DDQN-SS and PPO-SS [98]	Combines semantic segmentation with RL to provide scene-level understanding and improve perception and decision-making in complex road environments
Value-based MARL [103]	Enables cooperation between agents for lane keeping/changing by integrating fairness measurement and softmax Q-value calculation to ensure balanced and safe driving
DQN in RL-ACPAS [106]	Effectively learn and carry out collision avoidance techniques in real time, and improve the AVs' ability to make rational choices in unpredictable traffic conditions
SAC + feature extraction in Hybrid RL [107]	Provides effective complex sensory inputs process handling, and balancing exploration and exploitation by applying the motion-parameter adjustments
MARL + XAI [108]	Improve the decision-making module by integrating MARL and XAI, and increase the transparency and traceability of decisions made in AVs
GAIL-PPO [109]	Reduction of redundant input images and the high-quality data into decision-making module allow smoothness and traffic efficiency while maintaining vehicle safety
2-Stage Rule Learning [110]	Enhance the safety features as knowledge and intelligence aggregation is considered and transparency of decision can be improved
Adv. RAIM [111]	Provide complete collision elimination in the intersection scenario as the speed of each CAVs can be controlled through the proposed method
Dual-stream CNN + HFSM + DL [112]	Enable adaptability features in deciding the situation that encountered by the AVs and different method and strategy is used for simple and urban scenario

3. Analysis of Approaches

The decision-making development processes in AVs involve a structured breakdown of driving tasks. It starts with understanding the environment, proceeds to strategic interactions, and concludes with accurate motion control for AVs This section analyses the distinct roles and bridges them within four main decision-making tasks. In perception for situational awareness tasks, it is required to create a reliable and accurate description of the driving environment and surroundings for AVs, based on information collected by the allocated sensors. After that, the reviewed ML-based decision-making approaches, such as AdaBoost and SVM, will provide accurate and reliable decisions in structured environments. In the case of AdaBoost, it offers robust classification and prioritisation under time constraints, while SVM supports rapid lane-keeping and obstacle-avoidance decisions with optimisation techniques such as SSA or BOA.

Table 5. Comparative performance analysis of machine learning approaches categorised by decision-making tasks

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
Perception for Situational Awareness	AdaBoost [33] - SL	Improves decision-making robustness by combining weak classifiers for overtaking decisions.	Classification Accuracy	Accuracy = 87.64%	Real-world (Urban road)	• Limited to specific scenarios like cut-in manoeuvres • Struggles with uncertainty in real-world data
	SVM + FMCW [34] - SL	Classifies scenarios based on safety parameters for lane change and obstacle avoidance.	Classification Accuracy	Accuracy = 74% for the stationary and 33% for the moving target	Simulation (Urban)	• Narrows focus on lane changes • Lacks adaptability to complex traffic conditions
	Self-supervised learning + DSG strategy + Hybrid Losses [37] - HL	Eliminates the need for labelled depth data, enhances resilience by exposing the model to synthetic Gaussian noise, speckle noise, and dynamic blur and ensures stable depth predictions in dynamic scenes	RMSE (Root Mean Squared Error) Abs Rel (Absolute Relative Error)	Performance on clean data = same Accuracy on noisy data = 8.2%	Simulation	• Requires retraining for unseen disturbances, limiting its plug-and-play adaptability • Reliance on synthetic noise may not fully capture real-world complexities • The model could struggle with extreme motion blur or rapid scene changes
	Self-supervised learning + DNN [40] - HL	Enables training across large, diverse environments without manual labels, predicts occluded areas for safer path planning and better decision-making	Mean Intersection over Union (mIoU)	Best mIoU (FRRN-B) = 53.80% Baseline = 43.33%	Real-world (Rural and urban scenes)	• Performance depends on the quality of accumulated data • The model may struggle with highly dynamic scenes or rapid environmental changes, not well captured in training data
	Boosting + Clustering [40] - HL	Offers a reliable method for precisely identifying cars in a range of noise and environmental situations	Classification Accuracy Detection Time Detection Rate	Classification Accuracy = 98% Detection time shortened by 30% & False detection rate reduced by 20	Simulation (Urban)	• Heavily relies on the quality of the initial unsupervised clustering • Could misclassify or overlook vehicles due to limitations in the labelling process during clustering
	Transfer learning + Co-training + Data Augmentation [51] - HL	Improves robustness under adversarial conditions, generates high-confidence pseudo-labels from two different models and enhances the feature extraction quality from unlabelled data	Detection Accuracy	mAP@0.5 = ~89.6%	Simulation Env	• A wrong threshold leads to degraded results • Deteriorates in performance after co-training • The pipeline is complex and computationally intensive
	YOLOv8 [46] - DL	Provides safer navigation choices for AV in difficult real-world situations by improving object identification accuracy and dependability in a range of weather conditions	• Classification Accuracy • Precision, Recall Mean Average Precision (mAP)	• Precision = increment of 29.1% • Recall = 33% improvement mAP@.5 = 0.126 to 0.475	Real-world (Quebec City, Canada)	• Some object categories had very few instances, making the results statistically unreliable for these classes • The labelling process is time-consuming and requires significant human effort

Table 5. Continued

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
Perception for Situational Awareness	YOLOv8 [49] - DL	Provides high-accuracy, real-time identification of significant road objects, which provides a dependable perceptual basis for a driving support system to promptly warn of collisions.	• mean average precision (mAP)	• mAP = 88.2%	Real-world (roads from UMP Pekan campus to Pekan-Kuantan main road junction)	• Small Dataset • Limited Class Detection
	Data Fusion + Temporal awareness + RL in DriveSafeCore [56] - HL	Enhance the decision-making module by providing driving strategy that involve all the inputs from the perception module as the scene can be understood in depth through the multimodal sensor fusion and the help of the trajectory prediction, lead to the high accuracy in collision avoidance event and better safety decision during uncertainty scenario	• Accuracy rate • False positive rate • Inference latency • Throughout rate • Critical collision avoidance • Trajectory forecasting accuracy	• Accuracy rate = 98.56% • False positive rate = 1.23% • Inference latency = 42.3 ms • Throughout rate = 33.5 FPS • Critical collision avoidance = 96.78% • Trajectory forecasting accuracy = 96.1%	Real world scenarios (Static Obstacle Intrusion, Pedestrian Crossing, Highway Lane Cut-In, Sudden Stop of Lead Vehicle, Roundabout Navigation, Night-Time Driving, Foggy Weather Detection, Rainy Conditions)	• Required careful reward function designation • Depends on the accurate information from the perception module
Motion Planning and Trajectory Optimization	DQN + DDPG [32] - RL	Provides a robust, adaptive decision-making mechanism for autonomous vehicles and enables them to make informed decisions about context-aware driving decisions	• Travel time • Average acceleration • Average velocity	• Travel time = 94s (HDV = 149s) • Average velocity = 14.1 m/s (HDV = 8.5 m/s) • Average acceleration = 0.2 m/s² (HDV: 0.1 m/s²)	Simulation (three-lane traffic scene for 2 km)	• Needs a significant amount of interaction with the environment to learn effectively • Struggles with high-dimensional state and action spaces • Have difficulty in manual tuning • Limited adaptation in changing traffic conditions in real-time
	DL + Optimizer + MPC [84] - DL	Combines MPC and DNNs for optimised trajectory planning and collision avoidance	• Probability of Collision • Collision Cost	• Successful Collision Avoidance	Simulation (Two lane, Obstacles-rich)	• Restricts the model's ability to capture complex driving scenarios • Struggles to accurately classify smaller objects in images • Not adequately incorporate broader contextual information

Table 5. Continued

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
Motion Planning and Trajectory Optimization	POMDP (SARSOP, QMDP) + Decomposition method [85] - RL	Enhances the autonomous vehicle's capability to make safe and reliable decisions in uncertain environments with occluded users	• Collision rate • Time to cross (TTC)	• Collision Rate (Crosswalk) = 0.0% • Collision Rate (Intersection - QMDP) = 0.0% • Collision Rate (Intersection - SARSOP) = 0.7% • TTC (Crosswalk - SARSOP) = 10.51s • TTC (Crosswalk - QMDP) = 10.61s • TTC (Intersection – QMDP) = 6.20s • TTC (Intersection – SARSOP) = 4.38s	Simulated (Pedestrian crosswalk and intersection)	• Assumption of known users is not always feasible in real-time scenarios • Difficulties in generalisation across diverse driving situations. • Struggle in rapidly changing environments where user behaviours and surrounding conditions vary significantly
	POMDP + MCTS [90] - RL	Simulates human-like behaviour for collision prevention and overtaking, and handles uncertainty in decision-making for dynamic environments.	• Efficiency	• Probability Rate = 0.6	Simulation (Highway)	• Computationally complex • Struggles with real-time adaptation and model accuracy
	Homotopy boundary + A* planner [93] - DL	The classification of boundaries enable the safe gap identified between the ego and surrounding vehicles based on the expert demonstrations, helps in understanding the optimal maneuver strategy that should be executed to avoid collision	• Success Rate • Collision Rate • Timeout Rate	• Success Rate = 95.5% • Collision Rate = 4.0% higher • Timeout Rate = 0.5% lower	Simulation (Four-way unsignalised intersection scenarios)	• Only consider the longitudinal value • Required the expert demonstrations for the training purposes
	DL + IL + RL [94] - HL	Enable the high-level strategic decisions and adaptability features by ensembles the output of each learning-based for reliability enhancement, variance reduction and safety improvement	• Collision Rate • Route completion • Avg intervention frequency	• Collision Rate = 3.4% • Route completion = 97.2% • Avg intervention frequency = 0.4	Simulation (High-fidelity built-up surroundings with active weather and lighting situations)	• Increase the complexity of framework • Required the expert demonstrations for the training purposes

Table 5. Continued

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
Motion Planning and Trajectory Optimization	Transformer-transfer learning + uncertainty quantification [95] - HL	Proposed method enable the AVs to predict the trajectory of the HDVs accurately when perform lane change and by considering the uncertainty presence in the environment, potential collision between AVs and HDV can be avoided	• Average Distance Error (ADE) • Final Distance Error (FDE)	• ADE = 56.26% and 37.42% reduction (in Waymo dataset) • FDE = 69.60% and 56.86% reduction (in Waymo dataset) • ADE = 54.33% and 34.97% reduction (in driver simulator dataset) • FDE = 70.02% and 57.45% reduction (in driver simulator dataset)	Simulation (High-speed and low-speed scenario with the presence of AV and lane change of HDV)	• Dependency on the HDV's lane change dataset for training purpose • Required perfect detection information and high-quality interaction data
	MCTS + IL + NN [96] - HL	Known as transparent decision-making and safe planning as proposed method provide explicit exploration of action consequences over multiple time-steps and clear consequence of each nodes presented, lead to potential collision scenarios are identified and avoided	• Collision Rate (CR) • Driving Area Non-compliance (DA) • Ego Progress (EP)	• CR = 3% • DA = 2% • EP = 98%	Simulation Env	• Rely heavily on the model prediction accuracy • Sophisticated components lead to computational overhead
	Double-layer RL [97] - RL	Provide consistent optimization objectives based on the feedback provided by lower layer and trajectory-level reward function deployed in upper layer guide the parameter optimization of DRL decision-making model	• Safety Reward • Efficiency Reward • Comfort Reward • Decision Execution Reward	• Safety = -142.74 (two-lane) and -0.23 per timestep (three-lane) • Efficiency = -204.49 (two-lane) and -1.34 per timestep (three-lane) • Comfort = -10.72 (two-lane) and -0.04 per timestep (three-lane) • Decision Execution = -162.40 (two-lane) and -0.54 per timestep (three-lane)	Simulation (Multilane with the host vehicle and surrounding vehicles present)	• Rule-based constraint applied in the DRL model • Does not include the uncertainty element
	DDQN-SS + PPO-SS [98] - RL	Leverage semantic segmentation (DeepLabv3P) to enhance the vehicle's understanding of its environment, leading to more informed and stable decision - making processes in autonomous driving contexts	• Accuracy • mean intersection over union (MIOU)	• Driving stability (DQN-SS) = 14.2% • Driving stability (PPO-SS) = 28.5% • Reward increment = 25%	Simulation (CARLA clear noon conditions)	• A combined pipeline of semantic segmentation (SS) and reinforcement learning can introduce delays in decision-making • The SS network may not fully capture the complexities and variabilities of real-world driving conditions • Lower accuracy in SS can adversely affect the decision-making outcomes

Table 5. Continued

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
End-to-End and Multi-Agent Decision-Making	Fairness measurement & softmax operation in MARL [103] - RL	Ensures balanced performance between the cooperating vehicles, promoting cooperative behaviour while allowing for independent decision-making and improves the overall stability and performance of the learning process	• Average Speed • Cooperative Rate	• Average Speed (Heavy traffic) = 24.53 m/s • Average Speed (Loose Traffic) = 24.73 m/s • Cooperative Rate (Heavy Traffic) = 60.69 • Cooperative Rate (Loose Traffic) = 82.94	Simulation (Highly dynamic highway environments - Heavy Traffic & Loose Traffic)	• Reward function potentially leading to suboptimal decision-making in collaborative scenarios • Heavily relies on the quality and diversity of the training data • Leads to a substantial increase in the dimensionality of the state and action spaces
	DQN in RL-ACPAS [106] - RL	Enabling the system to learn optimal collision avoidance policies through continuous interaction	• Collision Avoidance Rate • Average Response Time • Efficiency Score	• Collision Avoidance Rate = >90% • Average Response Time: 50-70ms • Efficiency Score: 8.4 - 9.2	Simulation (Low Traffic Density, Moderate Traffic Density, High Traffic Density)	• Simulation-based training
	SAC + feature extraction in HRL [107] - RL	Able to handle and process the complex sensory inputs from radar and LiDAR effectively and maximise cumulative reward while maintaining policy entropy	• Collision avoidance reward value • Collision Avoidance Rate (CAR) • Average Reaction Time (ART)	• Collision avoidance reward = Higher • CAR = High success rate • ART = Low	Simulation and real scene	• Requires a large amount of computing power • The real-world environment is not fully represented by model-based components
	MARL + XAI [108] - RL	Integration of MARL and XAI increases the robustness of the decision-making framework in terms of fewer collisions in the simulated environment and a higher score of SHAP	• Success Rate • Decision Latency • Collision Rate • SHAP Score	• Success Rate = 83% • Decision Latency = 132 ms • Collision Rate = 4.6 / 100 km • SHAP Score = 0.77	Simulation (Ambiguous urban scenarios)	• Requires human-supervision module to merge the decisions from each agent in MARL • Low simulated episodes (200 episodes)
	GAIL-PPO [109] - HL	Utilisation of attention mechanism and inter-frame redundancy reduction strategy in providing high quality input information to the decision-making module allow improvement and robustness in decision made as well as safety enhancement	• Learning efficiency • Safety • Smoothness • Efficiency	• Learning efficiency = 28.15% • Safety = 30.28% • Smoothness = 5.46% • Efficiency = 7.8%	Simulation (Urban and street scenario)	• Reliance on the experts prior knowledge for RL model • Reduction on the decision stability
	2-Stage Rule Learning [110] - HL	Enhanced safety of driving and collision reduction in CAVs as the knowledge and intelligence aggregation enable continuous learning in changing environment	• Prediction accuracy • Decision transparency	• Prediction accuracy = detect more than 90% unsafe state transitions • Decision transparency = outperforms the VFDT baseline	Simulation (Highway driving)	• Potential for False Positives in Unsafe Predictions • Limited Adaptability to Complex Unstructured Driving Environments

Table 5. Continued

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
End-to-End and Multi-Agent Decision-Making	Adv. RAIM [111] - RL	Enable the provided decision is collision free as the state/conflict encoder module deployed, ensure the environment is understood and features of each CAVs is fed into input	• Travel time • Time lost due to congestion • Waiting time	• Travel time = 59% reduction • Time lost due to congestion = 95% reduction • Waiting time = 56% reduction	Simulation (Intersection)	• Required data increment once the solution performance has maxed • Need extensive training time
	Dual stream CNN + HFSM + DL [112] - HL	Proposed hybrid decision-making framework enable the dynamic and adaptive strategy selection based on the processed input from the dual-stream CNN, lead to safety enhancement and collision avoidance in diverse environment using HFSM for simple environment and DL for complex environment	• Decision optimality • Safety rate • Efficiency • Lane idleness rate	• Decision optimality = 30% improvement • Safety rate = 92% • Efficiency = 66% • Lane idleness rate = 85.5%	Simulation (Town06 - Highway, Town02 - Urban)	• DL required high computational resources • Do not focus on the safety feature
Behavioural-Level Decision-Making	AdaBoost [57] - SL	Improves decision-making robustness by combining weak classifiers for overtaking decisions.	• Classification Accuracy	• Accuracy = >90% in 1s, 2s and 3s time constraint • Validation accuracy = ~0.8	Simulation (Countryside road + four lanes + pedestrians crossing)	• Limited to specific scenarios like cut-in manoeuvres • Struggles with uncertainty in real-world data
	SVM + SSA [61] - SL	Classifies scenarios based on safety parameters for lane change and obstacle avoidance.	• Classification Accuracy	• Accuracy rate = 93.6%	Simulation (Roadway)	• Narrow focus on lane changes • Lacks adaptability to complex traffic conditions
	SVM + BOA [62] - SL	Classifies scenarios based on safety parameters for lane change and obstacle avoidance.	• Classification Accuracy	• Accuracy = >85%	Simulation (Training & Validation) & Real-vehicle Zhongtong bus)	• Narrows focus on lane changes • Lacks adaptability to complex traffic conditions
	Standard SVM [64] - SL	Classifies scenarios based on safety parameters for lane change and obstacle avoidance.	• Classification Accuracy	• Accuracy = 94.2%	Simulation	• Narrows focus on lane changes • Lacks adaptability to complex traffic conditions
	Rule-based + Supervised learning [65] - SL	Analyses real-time vehicle states, environment, and traffic data, refines the decision model's outputs by learning optimal braking and steering scale coefficients from previous control performance and real-time sensor input	• Vehicle Stability • Collision Avoidance Effectiveness	• Low-friction road: performance improvement = ~25% • High Friction Road: performance improvement = ~24.3%	Simulation (Two lane with right lane occupied with jamming vehicle and on left lane, obstacle in front of host vehicle)	• The effectiveness of the self-learning framework heavily depends on an accurate perception of traffic scenario information through sensors • Usage of multilayer neural networks and dynamic decision models could introduce computational challenges

Table 5. Continued

Decision-making Task	Approaches & ML category	Robustness of Decision-Making Algorithms	Key Metrics	Reported Performance	Test Environment	Limitation
Behavioural-Level Decision-Making	Policy search + Self-Supervised learning [68] - HL	Generate collision-free, smooth lane change manoeuvres, enhancing its robustness in varied traffic scenarios and providing a more resilient decision-making process	• Collision Avoidance Rate • Manoeuvre Optimality	• Rate = 104% • Manoeuvre Optimality = Successful convergence of MPC optimization problem	Simulation (Two lanes, with one lane with ego vehicle and another lane with target lane)	• Inaccuracies in models could compromise the effectiveness of the generated manoeuvres • Could impose computational burdens in real-time applications
	LSTM [71] - DL	Predicts lane change intentions by analysing temporal dependencies in driving behaviour	• Classification Accuracy • Mean Square Error (MSE)	• Accuracy rate = >93.0% • MSE = >0.0949	Simulation (Two-way four lane expressway)	• Requires extensive training data • Struggles with dynamic environmental changes
	Linear algebra-based [79] - DL	Facilitates computational efficiency and enhances the clarity and reliability of decisions made by autonomous vehicles	• Mean pixel accuracy (mPA) • Mean intersection over the union (mIoU)	• mPA = 92.95% • mIoU = 86.82%	Simulation	• Computationally intensive • Struggles with real-time decision-making in complex environments
	Fuzzy C-means + XGBoost + RFE [80] - HL	Offer behaviour assessment and predict the risk levels of instantaneous next-step movements	Classification Accuracy	• Overall accuracy predictive power = 91.7% • Safe-risk distinction accuracy = ~95%	Simulation	• Lead to diminishing returns due to noisy data and unmeasured hidden variables • Lack of real-world crash instances in the raw dataset
	DQN + DDPG [32] - RL	Provides a robust, adaptive decision-making mechanism for autonomous vehicles and enabling them to make informed about context-aware driving decisions	• Travel time • Average acceleration • Average velocity	• Travel time = 94s (HDV = 149s) • Average velocity = 14.1 m/s (HDV = 8.5 m/s) • Average acceleration = 0.2 m/s² (HDV: 0.1 m/s²)	Simulation (three-lane traffic scene for 2 km)	• Need a significant amount of interaction with the environment to learn effectively • Struggle with high-dimensional state and action spaces • Have difficulty in manual tuning • Limited adaptation in changing traffic conditions in real-time
	GMA-TD3 [82] - RL	Provide effective sequential avoidance decision in the encountered scenarios based on the detected motion and utilized the dual-level reward mechanism	• Navigation task completion time • Path length in challenging scenarios • Reward convergence speed	• Navigation task completion time = least timesteps (490-578) • Path length = shortest (13.63 n mile – 16.06 n mile)	Simulation (3 scenarios: Multi-ship, Two-ship and encounter for static and dynamic obstacles)	• Unpredictable behaviour of conventional obstacles • Utilized single RL agent • Required tracking accuracy enhancement

For DL of YOLOv8, it is used to identify objects in real time, and self-supervised learning is deployed to estimate environmental depth. The combination of 2 methods yields robustness to adversarial attacks and occlusions, which is crucial for real-world AV safety. In this task, the focus has shifted from detection accuracy to robustness to distribution shifts. It also involves handling noise, occlusion, and diverse environmental conditions. For the behavioural-level decision-making task, the tactical and intent-driven decisions are generated based on the sensory output from the perception task. The utilised approaches are varied, which reflects the wide range of driving manoeuvres involved. For example, the SL approach can successfully derive explicit policies from human imitation, while HL enables the merging of policy search with self-supervised learning to provide more adaptive manoeuvre development. Meanwhile, the implementation of RL led to a major accomplishment: DQN. The DQN agent interacts with the environment directly to acquire the lane-change technique. This indicates the transition from imitation-based learning to optimisation-driven behaviour. These activities are adjusted according to reward functions that prioritise stability, safety, and efficiency in AVs However, the integration technique and the exploratory nature of RL may pose a major problem for safety criteria. Thus, it is important to ensure that the generated policies remain both efficient and demonstrably safe in previously unexpected driving scenarios.

As for the motion planning and trajectory optimisation task, it converts the decisions from the behavioural-level decision-making task into real actions. Examples of real actions that will be executed by AVs are lane change, lane keep, accelerate, decelerate, steer, and brake. These actions enable collision-avoidance events in AVs While considering obstacles, boundaries, traffic laws, and vehicle dynamics, the task will address continuous optimisation issues in AVs Currently, most approaches rely on sophisticated RL and model-based optimisation. The MPC promotes a well-established method for limited trajectory optimisation. However, its efficiency depends significantly on reliable systems and prediction models. Hybrid DL–MPC techniques leverage deep neural networks to predict complex dynamics or collision-related cost functions, addressing modelling limitations. In cases of uncertainty, frameworks based on POMDPs are combined with solvers such as SARSOP or online techniques like MCTS to provide theoretically principled solutions. Besides, RL algorithms, such as DDPG, are used to learn continuous control policies in car-following scenarios. The primary difficulty lies in finding a way to maintain both real-time computational speed and the high accuracy needed to simulate complex nonlinear dynamics and unpredictable environmental factors. As for end-to-end and multi-agent decision-making tasks, it usually combines the classic modular pipeline to tackle system-wide consistency and coordination among multiple actors. The goals of the multi-agent and end-to-end decision-making tasks are to establish balanced outcomes in shared settings (Multi-agent) and to master direct pixel-to-control mappings (End-to-End). An example of an end-to-end approach is DDQN-SS. The approach combines perception and control into a single network and then uses it to optimise the global driving objective. Even though the method offers greater flexibility, it still has disadvantages regarding safety verification and interpretability. In MARL, the decision-making task influences the modelling of other agents' policies, fostering collaboration and ensuring fairness. The techniques, such as fairness measurements and softmax operations in value networks, are designed to stabilise learning and foster equitable cooperation among agents. However, the main obstacle is the non-stationarity of the learning environment. This occurs when all agents must learn simultaneously, violating the fundamental premise of a stationary MDP. It results in safety guarantees and convergence difficulty.

In the ML-based context, each approach has limitations. For SL, these models tend to struggle in unstructured or unpredictable driving conditions because they rely on fixed data patterns. Meanwhile, for HL, hybrid models always increase computational complexity, make pseudo-label thresholds more sensitive, and increase dependence on simulation-based or synthetic data. This will affect scalability and real-world reliability. As for DL, the constraints include the need for extensive labelled data for training and the computational intensity of training. These constraints will limit the DL strength in real-world deployment. RL faces challenges such as high training costs, sample inefficiency, and real-time scalability. Beyond these technical constraints, ethical decision-making is a major non-technical barrier across all ML categories, especially in situations where collisions are unavoidable. Notably, the ethical decision-making system proposed for AVs reduces the loss of life by an average of 24.11% compared to the standard algorithm [113]. Even though various supervised techniques, such as AdaBoost [57], have been used to address this issue, there is no standard framework in the field to address this ethical decision-making challenge. Therefore, a crucial and unexplored field of research is required to incorporate ethical reasoning to balance occupant and pedestrian safety. Adjustments to traffic laws also contribute to lowering unavoidable collisions by implementing the cost functions of hybrid models or the reward functions of RL in AVs. This challenge highlights the requirement of improving algorithmic performance and value-sensitive design to produce safe and socially acceptable AVs. Table 5 presents a comparative performance analysis of ML approaches categorised by decision-making tasks used to address the significant challenge of collision avoidance in AVs. The table is structured based on decision-making tasks and ML algorithmic approaches. Table 5 presents the progression of tasks, from simple perception, such as situational awareness, to more complex sequential decision-making, including behavioural-level decision-making and motion planning. In the early-stage tasks, the SL and HL approaches are utilised for classification and prediction, using methods such as AdaBoost for overtaking and YOLOv8 for object detection. In contrast, for later-stage tasks, it focuses on planning and control in AVs, which mainly use the RL and DL frameworks such as DQN, DDPG, and POMDP. This indicates a shift towards learning optimal action policies in dynamic and unpredictable environments.

4. Conclusions

In recent years, machine learning has emerged as a key component of intelligent decision-making in AVs, especially for collision avoidance. In this review, ML-based decision-making approaches are analysed and systematically categorised into four major branches: supervised learning, hybrid learning, deep learning, and reinforcement learning. The applicability of each approach presented is evaluated based on 4 key autonomous driving tasks: perception for situational awareness, behavioural-level decision-making, motion planning, and multi-agent coordination. This structured classification provides a comprehensive view of how various ML categories improve the effectiveness and safety of AVs. Supervised learning methods such as AdaBoost and SVMs offer reliable performance in structured environments and are highly effective for tasks like lane-changing and vehicle behaviour classification. However, they also come with limitations. Their applicability in complex and uncertain environments is limited because supervised learning models still rely on labelled data and are poorly adaptable to dynamic contexts. As for hybrid learning models, they improve upon this issue by integrating unsupervised components. This results in prediction accuracy enhancement and feature selection through combinations such as Fuzzy C-Means with XGBoost. Even though the hybrid learning model is efficient in theory, it faces practical challenges, including generalisation issues and high computational costs.

Deep learning approaches, including LSTM networks, CNNs, and integrated architectures such as Deep Learning-based MPC, demonstrate strong capabilities in perception, temporal reasoning, and planning. Nevertheless, they have limitations, such as high data requirements and computational demands. These deep learning models also have limited robustness in novel or ethically ambiguous scenarios. In comparison with reinforcement learning (RL), these RL models benefit from a highly flexible and scalable framework for autonomous systems. The RL agents are required to learn directly through interaction with the environment and to allow for dynamic adaptation. Recent advancements in RL have produced sophisticated frameworks such as DDQN-SS and MARL. These sophisticated RL models have improved trajectory prediction and cooperative behaviours in complex traffic scenarios. Furthermore, the deployed enhancements in RL agents, such as DDPG, are integrated with Gaussian Processes and planners based on POMDPs, demonstrating that decision-making under partial observability can be improved. Despite these enhancements, RL approaches still face significant challenges, such as sample inefficiency, long training times, and substantial computational demands that delay real-time implementation.

Based on the comparative analysis shown in Table 5, no single ML approach emerges as a universal solution for collision avoidance in autonomous driving for AVs. Each approach exhibits distinct strengths and limitations, whether in data dependence, computational load, or handling uncertainty. As a result, future decision-making systems in AVs are required to prioritise the development of hybrid or context-aware frameworks that integrate the complementary strengths of approaches across these methodologies. For example, when RL's adaptability is integrated with model-based controllers, such as MPC, this integration can increase safety guarantees or strengthen ethical reasoning within the AVs' cost and reward functions. Another thing is that the field must progress towards real-world validation. Even though the designed framework has been tested in simulations and provides valuable insights, it still falls short of replicating the intricate and unpredictable environment of actual traffic conditions. Besides that, future research should focus on developing lightweight, edge-deployable models, advancing probabilistic reasoning under uncertainty, and enhancing the integration of ethical frameworks within AV decision modules. Additionally, the deployment of multi-agent systems will support coordinated behaviours and intention prediction in connected environments, as it holds significant promise for improving overall traffic safety and efficiency. Overall, this review contributes a structured classification and comparative synthesis of ML-based approaches for AVs' collision avoidance by bridging the perception, decision-making, and control modules. It benefits researchers and practitioners by providing a specific roadmap towards safer, more robust, and socially acceptable autonomous driving systems, emphasising task-specific capabilities, limitations, and interdisciplinary gaps.

Acknowledgements

The authors would like to thank Universiti Teknologi MARA (UiTM) for supporting this research project.

Funding

The authors would like to thank Universiti Teknologi MARA (UiTM) for supporting this research project through Geran Insentif Penyelidikan (GIP) 600-RMC/GIP 5/3 (029/2024).

Declaration of Competing Interests

The authors declare no conflicts of interest.

CRediT Authorship Contribution Statement

Siti Nurhafizza Maidin: Conceptualisation; Methodology; Validation; Analysis; Writing – original draft; Supervisor
Nurbaiti Wahid: Conceptualisation; Methodology; Validation; Analysis; Writing – revise & review; Supervisor
Sarah Atifah Binti Saruchi: Analysis; Writing – revise & review
Sumiya Tamanna Fujita: Analysis; Writing – revise & review
Alvi Khan Chowdhury: Writing – revise & review

Availability of Data and Materials

The data supporting this study's findings are available on request from the corresponding author.

Ethics Declarations
This study did not involve human participants or animals. Ethical approval was therefore not required.

Generative Artificial Intelligence Declarations
The authors claim that artificially intelligent-assisted technologies, such as generative AI, were not used to generate content, ideas, or theories. We have just utilised AI to enhance readability and refine the language. This was used with extreme human control and oversight. The authors take full responsibility for reviewing and approving the content.

References
[1] M. Faudzi and M. Yusof, "Malaysian value of statistical life for fatal injury in road accident: A conjoint analysis approach," *Journal of Society for Transportation and Traffic Studies*, vol. 2, no. 2, pp. 30–40, 2004.
[2] M. H. Azmi, "Statistik kemalangan, kemalangan maut dan kematian jalan raya mengikut jenis jalan," Sep. 2021. https://www.data.gov.my/data/ms_MY/dataset/statistik-laporan-kemalangan-jalan-raya-diterima-2010-2017 (accessed Dec. 07, 2022).
[3] Autonomous Vehicle Industry Association (AVIA), "What is an autonomous vehicle," Jan. 30, 2023. https://theavindustry.org/resources/fact-sheets/pdf/What-is-an-AV.png (accessed Dec. 23, 2024).
[4] J. Wang, L. Zhang, Y. Huang, J. Zhao, and F. Bella, "Safety of autonomous vehicles," *Journal of Advanced Transportation*, vol. 2020, pp. 1–13, 2020. https://doi.org/10.1155/2020/8867757
[5] C. Badue, R. Guidolini, R.V. Carneiro et al., "Self-driving cars: A survey," *Expert Systems with Applications*, vol. 165, p. 113816, 2021. https://doi.org/10.1016/j.eswa.2020.113816
[6] D. Garikapati and S. S. Shetiya, "Autonomous vehicles: Evolution of artificial intelligence and the current industry landscape," *Big Data and Cognitive Computing*, vol. 8, no. 4. p. 42, Apr. 07, 2024. https://doi.org/10.3390/bdcc8040042
[7] J. Wu, J. Zhang, B. Nie, Y. Liu, and X. He, "Adaptive control of PMSM servo system for steering-by-wire system with disturbances observation," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 2, pp. 2015–2028, 2022. https://doi.org/10.1109/TTE.2021.3128429
[8] J. Zhang, J. Wu, J. Liu et al., "Command-filter-adaptive-based lateral motion control for autonomous vehicle," *Control Engineering Practice*, vol. 121, p. 105044, 2022. https://doi.org/10.1016/j.conengprac.2021.105044
[9] D. González, J. Pérez, V. Milanés, and F. Nashashibi, "A review of motion planning techniques for automated vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 1135–1145, 2016. https://doi.org/10.1109/TITS.2015.2498841
[10] H. Fan, F. Zhu, C. Liu, et al., "Baidu Apollo EM motion planner," *arXiv preprint arXiv:1807.08048*, 2018. [Online]. Available: http://arxiv.org/abs/1807.08048
[11] C. Liu, C. Y. Lin, and M. Tomizuka, "The convex feasible set algorithm for real time optimization in motion planning," *SIAM Journal on Control and Optimization*, vol. 56, no. 4, pp. 2712–2733, 2018. https://doi.org/10.1137/16M1091460
[12] S. M. Veres, L. Molnar, N. K. Lincoln, and C. P. Morice, "Autonomous vehicle control systems — a review of decision making," in *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 225, no. 2, pp. 155–195, 2011. https://doi.org/10.1177/2041304110394727
[13] M. F. Yasak, P. M. Heerwan, and V. R. Aparow, "Collision avoidance strategies in autonomous vehicles and on-ramp scenario: A review," *Annual Reviews in Control*, vol. 59, no. February, p. 100986, 2025. https://doi.org/10.1016/j.arcontrol.2025.100986
[14] J. Hardy and M. Campbell, "Contingency planning over probabilistic obstacle predictions for autonomous road vehicles," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp. 913–929, 2013. https://doi.org/10.1109/TRO.2013.2254033
[15] Z. Li, C. Lu, Y. Yi, and J. Gong, "A hierarchical framework for interactive behaviour prediction of heterogeneous traffic participants based on graph neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9102–9114, 2022. https://doi.org/10.1109/TITS.2021.3090851
[16] Z. Li, J. Gong, C. Lu, and Y. Yi, "Interactive behavior prediction for heterogeneous traffic participants in the urban road: a graph-neural-network-based multitask learning framework," *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 3, pp. 1339–1349, 2021. https://doi.org/10.1109/TMECH.2021.3073736
[17] P. Fiorini and Z. Shiller, "Motion planning in dynamic environments using velocity obstacles," *International Journal of Robotics Research*, vol. 17, no. 7, pp. 760–772, 1998. https://doi.org/10.1177/027836499801700706
[18] A. J. M. Muzahid, S. F. Kamarulzaman, M. A. Rahman, S. A. Murad, M. A. S. Kamal, and A. H. Alenezi, "Multiple vehicle cooperation and collision avoidance in automated vehicles: survey and an AI-enabled conceptual framework," *Scientific Reports*, vol. 13, no. 1, p. 603, 2023. https://doi.org/10.1038/s41598-022-27026-9
[19] R. Tami, B. Soualmi, A. Doufene, J. Ibanez, and J. Dauwels, "Machine learning method to ensure robust decision-making of AVs," in *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, Oct. 2019, pp. 1217–1222. https://doi.org/10.1109/ITSC.2019.8917085
[20] G. De-Las-Heras, J. Sánchez-Soriano, and E. Puertas, "Advanced driver assistance systems (ADAS) based on machine learning techniques for the detection and transcription of variable message signs on roads," *Sensors*, vol. 21, no. 17, pp. 1–18, 2021. https://doi.org/10.3390/s21175866
[21] N. E. Titus, "The role of advanced driver assistance systems (ADAS) In improving road safety," vol. 3, no. 2, pp.

2–5, 2024.
[22] H. Sun, T. Tanprasert, and R. Setthawong, “A collaborative study of perception and decision-making using machine learning in autonomous vehicles,” *Journal of Logistics, Informatics and Service Science*, vol. 11, no. 2, pp. 433–450, 2024. https://doi.org/10.33168/JLISS.2024.0227
[23] V. Goar and N. S. Yadav, *Foundations of machine learning*. 2024. https://doi.org/10.4018/979-8-3693-1598-9.ch002
[24] Y. Baştanlar and M. Ozuysal, *Introduction to machine learning Second Edition*, vol. 1107, 2014. https://doi.org/10.1007/978-1-62703-748-8_7
[25] M. Kubat, *An Introduction to Machine Learning*, 2017. https://doi.org/10.1007/978-3-319-63913-0
[26] I. H. Sarker, “Deep Learning: A comprehensive overview on techniques, taxonomy, applications and research directions,” *SN Computer Science*, vol. 2, no. 6, pp. 1–20, 2021. https://doi.org/10.1007/s42979-021-00815-1
[27] R. K. Mishra, G. Y. S. Reddy, and H. Pathak, “The understanding of deep learning: A comprehensive review,” *Mathematical Problems in Engineering*, vol. 2021, p. 5548884, 2021. https://doi.org/10.1155/2021/5548884
[28] M. F. Seldin, R. Shigeta, P. Villoslada et al., “European population substructure: Clustering of northern and southern populations,” *PLoS Genetics*, vol. 2, no. 9, pp. 1339–1351, 2006. https://doi.org/10.1371/journal.pgen.0020143
[29] I. Basu, S. Karmakar, S. K. Kundu, A. Saha, and G. S. Taki, “Application of reinforcement learning for control of autonomous vehicles,” in *2022 International Conference on Interdisciplinary Research in Technology and Management, IRTM 2022 - Proceedings*, 2022. https://doi.org/10.1109/IRTM54583.2022.9791531
[30] Z. Yang, X. Pei, J. Xu, X. Zhang, and W. Xi, “Decision-making in Autonomous driving by reinforcement learning combined with Planning & control,” in *2022 6th CAA International Conference on Vehicular Control and Intelligence, CVCI 2022*, Oct. 2022, pp. 1–6. https://doi.org/10.1109/CVCI56766.2022.9964691
[31] B. R. Kiran, I. Sobh, V. Talpaert et al., “Deep reinforcement learning for autonomous driving: A survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2022. https://doi.org/10.1109/TITS.2021.3054625
[32] Y. Shi, J. Liu, C. Liu, and Z. Gu, “DeepAD: An integrated decision-making framework for intelligent autonomous driving,” *Transportation Research Part A: Policy and Practice*, vol. 183, p. 104069, 2024. https://doi.org/10.1016/j.tra.2024.104069
[33] X. Long, H. Ming, and T. Tao, “Vehicle tracking based on Adaboost cascade classifier,” *Journal of Physics: Conference Series*, vol. 1650, no. 3, p. 032118, 2020. https://doi.org/10.1088/1742-6596/1650/3/032118.
[34] S. Chaudhary, A. Sharma, S. Khichar, Y. Meng, and J. Malhotra, “Enhancing autonomous vehicle navigation using SVM-based multi-target detection with photonic radar in complex traffic scenarios,” *Scientific Reports*, vol. 14, no. 1, p. 17339, 2024. https://doi.org/10.1038/s41598-024-66850-z
[35] J. Tebart, M. Steeg, and A. Stohr, “Photonics-based FMCW radar localization using direct laser modulation and leaky-wave antenna beam scanning,” *International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz*, vol. 2020-Novem, pp. 326–327, 2020. https://doi.org/10.1109/IRMMW-THz46771.2020.9370727
[36] R. Dhawan, K. Moyal, and A. Choudhary, “Optical filter-less photonic FMCW radar for multi-target detection,” *IEEE Photonics Technology Letters*, vol. 35, no. 2, pp. 81–84, 2023. https://doi.org/10.1109/LPT.2022.3224219
[37] C. Zhao, J. Zhao, Y. Xu, and J. An, “DSG:A Robustness-enhanced self-supervised depth estimation method,” *27th IEEE/ACIS International Summer Conference on Software Engineering Artificial Intelligence Networking and Parallel/Distributed Computing, SNPD 2024 - Proceedings*, pp. 91–96, 2024. https://doi.org/10.1109/SNPD61259.2024.10673955
[38] E. Wang, Zhou; Bovik, Alan Conrad; Rahim Sheikh, Hamid; P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004, https://doi.org/10.1109/TIP.2003.819861
[39] A. K. Venkataramanan, C. Wu, A. C. Bovik, I. Katsavounidis, and Z. Shahid, “A hitchhiker’s guide to structural similarity,” *IEEE Access*, vol. 9, pp. 28872–28896, 2021. https://doi.org/10.1109/ACCESS.2021.3056504
[40] J. Stojcheski, T. Nurnberg, M. Ulrich, T. Michalke, C. Glaser, and A. Geiger, “Self-supervised occupancy grid map completion for automated driving,” *IEEE Intelligent Vehicles Symposium, Proceedings*, vol. 2023-June, 2023, https://doi.org/10.1109/IV55152.2023.10186748
[41] T. Collins, J. J. Collins, and C. Ryan, “Occupancy grid mapping: An empirical evaluation,” *2007 Mediterranean Conference on Control and Automation, MED*, 2007. https://doi.org/10.1109/MED.2007.4433772
[42] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017. [Online]. Available: https://arxiv.org/pdf/1511.00561.pdf
[43] W. Weng and X. Zhu, “U-Net: convolutional networks for biomedical image segmentation,” *arXiv*, vol. 9, pp. 16591–16603, 2015. https://doi.org/10.1109/ACCESS.2021.3053408
[44] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, “U-Net and its variants for medical image segmentation: a review of theory and applications,” *IEEE Access*, vol. 9, pp. 82031–82057, 2021. https://doi.org/10.1109/ACCESS.2021.3086020
[45] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe, “Full-resolution residual networks for semantic segmentation in street scenes,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 3309–3318, 2017. https://doi.org/10.1109/CVPR.2017.353

[46] T. Sharma, A. Chehri, I. Fofana, B. Debaque, N. Duclos, and S. Khare, "Deep learning-based road object detection for collision avoidance in autonomous driving," *Proceedings - 2024 IEEE World Forum on Public Safety Technology, WFPST 2024*, no. May, pp. 126–131, 2024. https://doi.org/10.1109/WFPST58552.2024.00020
[47] R. Varghese and M. Sambath, "YOLOv8: A novel object detection algorithm with enhanced performance and robustness," *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems, ADICS 2024*, pp. 1–6, 2024. https://doi.org/10.1109/ADICS58448.2024.10533619
[48] Vivi Afifah and S. Erniwati, "YOLOv8 for object detection: a comprehensive review of advances,techniques, and applications," *IJACI : International Journal of Advanced Computing and Informatics*, vol. 2, no. 1, pp. 53–61, 2025. https://doi.org/10.71129/ijaci.v2i1.pp53-61
[49] M. S. Beg, M. Y. Ismail, M. Saef Ullah Miah, and M. H. Peeie, "Enhancing driving assistance system with YOLO V8-Based normal visual camera sensor," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 31, no. 1, pp. 226–236, 2023. https://doi.org/10.37934/ARASET.31.1.226236
[50] J. Duan, H. Ye, H. Zhao, and Z. Li , "Deep cascade adaboost with unsupervised clustering in autonomous vehicles," *Electronics (Switzerland)*, vol. 12, no. 1, p. 44, 2023. https://doi.org/10.3390/electronics12010044
[51] W. Chen, J. Yan, W. Huang, W. Ge, H. Liu, and H. Yin, "Robust object detection for autonomous driving based on semi-supervised learning," *Security and Safety*, vol. 3, p. 2024002, 2024. https://doi.org/10.1051/sands/2024002
[52] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," *arXiv*, pp. 9729–9738, 2019.
[53] X. Chen, H. Fan, R. Girshick, and K. He, "Improved baselines with momentum contrastive learning," *arXiv*, pp. 1–3, 2020. [Online]. Available: http://arxiv.org/abs/2003.04297
[54] X. Liu, H. Yang, Z. Liu, L. Song, H. Li, and Y. Chen, "DPatch: An adversarial patch attack on object detectors," *arXiv preprint arXiv:1806.02299*, 2019.
[55] A. Saha, A. Subramanya, K. Patil, and H. Pirsiavash, "Role of spatial context in adversarial robustness for object detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2020-June, pp. 3403–3412, 2020. https://doi.org/10.1109/CVPRW50498.2020.00400.
[56] K. L. Prabha, J. G. Thatipudi, M. Satyavathi, M. Vohra, B. Viswanathan, and S. Bhattacharjee, "DriveSafeCore: Real-time collision avoidance through spatio-temporal deep learning in autonomous vehicles," *2025 6th International Conference on Data Intelligence and Cognitive Informatics, ICDICI 2025*, pp. 1431–1438, 2025. https://doi.org/10.1109/ICDICI66477.2025.11135181.
[57] A. Singh, Y. Murzello, S. Pokhrel, and S. Samuel, "Classification of moral decision making in autonomous driving: efficacy of boosting procedures," *Information*, vol. 15, no. 9, p. 562, 2024. https://doi.org/10.3390/info15090562.
[58] R. E. Schapire, "A brief introduction to boosting," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2, pp. 1401–1406, 1999.
[59] R. E. Schapire, "The boosting approach to machine learning: An Overview BT - nonlinear estimation and classification," *Nonlinear Estimation and Classification*, vol. 171, no. Chapter 9, pp. 149–171, 2003. https://doi.org/doi/10.1007/978-0-387-21579-2_9
[60] M. Stamp, "Boost your knowledge of AdaBoost," no. October, 2019.
[61] J. Wang, H. Wang, M. Fei, and G. Zhou, "Vehicle lane changing game model based on improved SVM algorithm," *World Electric Vehicle Journal*, vol. 15, no. 11, 2024. https://doi.org/10.3390/wevj15110505.
[62] Y. Liu, X. Wang, L. Li, S. Cheng, and Z. Chen, "A novel lane change decision-making model of autonomous vehicle based on support vector machine," *IEEE Access*, vol. 7, pp. 26543–26550, 2019. https://doi.org/10.1109/ACCESS.2019.2900416.
[63] X. Wang, J. Fan, and N. Liu, "A novel decision-making algorithm of autonomous vehicle based on improved SVM," in *ACM International Conference Proceeding Series*, 2020, pp. 643–648. https://doi.org/10.1145/3434581.3434706
[64] M. Karthikeyan, S. Sathiamoorthy, and M. Vasudevan, "Lane keep assist system for an autonomous vehicle using support vector machine learning algorithm," *Lecture Notes on Data Engineering and Communications Technologies*, vol. 46, no. December, pp. 101–108, 2020. https://doi.org/10.1007/978-3-030-38040-3_11
[65] Y. Wang, G. Yin, Y. Li et al., "Self-learning control for coordinated collision avoidance of automated vehicles," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 235, no. 4, pp. 1149–1163, 2021. https://doi.org/10.1177/0954407019887884
[66] A. Kane and A. Hussain, "Artificial neural networks - An overview," *Indian Drugs*, vol. 30, no. 5, pp. 168–178, 1993. https://doi.org/10.58496/mjcsc/2023/015
[67] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, no. 1, pp. 3–31, 2000. https://doi.org/10.1016/S0167-7012(00)00201-3
[68] F. Laneve, A. Rucco, and M. Bertozzi, "A hybrid approach combining upper-level policy search and mpc for lane change of autonomous vehicles," *2024 European Control Conference, ECC 2024*, pp. 3156–3163, 2024. https://doi.org/10.23919/ECC64448.2024.10590819
[69] R. A. Levine and G. Casella, "Implementations of the monte carlo EM algorithm," *Journal of Computational and Graphical Statistics*, vol. 10, no. 3, pp. 422–439, 2001. https://doi.org/10.1198/106186001317115045
[70] W. Ruth, "A review of Monte Carlo-based versions of the EM algorithm," *arXiv*, 2024, [Online]. Available: https://arxiv.org/abs/2401.00945v1

[71] T. Wang, Z. Guan, T. Peng, D. Li, and R. Zhao, "Human-like Decision Making for Autonomous Lane Changing Using Deep Learning," in *Proceedings of 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence, ICIBA 2021*, 2021, pp. 950–955. https://doi.org/10.1109/ICIBA52610.2021.9688311

[72] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[73] R. C. Staudemeyer and E. R. Morris, "Understanding LSTM -- a tutorial into long short-term memory recurrent neural networks," pp. 1–42, 2019. [Online]. Available: http://arxiv.org/abs/1909.09586

[74] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15, 2015.

[75] J. Fan, Z. Wang, Y. Xie, and Z. Yang, "A Theoretical analysis of deep Q-learning," *Proceedings of Machine Learning Research*, vol. 120, pp. 486–489, 2020.

[76] M. Sewak, *Deep Reinforcement Learning*, Singapore: Springer Nature Singapore Pte Ltd., 2019. https://doi.org/10.1007/978-981-13-8285-7

[77] E. H. Sumiea, S.J. Abdulkadir, H.S. Alhussian et al., "Deep deterministic policy gradient algorithm: A systematic review," *Heliyon*, vol. 10, no. 9, p. e30697, 2024. https://doi.org/10.1016/j.heliyon.2024.e30697

[78] V. Sumalatha and S. Pabboju, "An overview of deep deterministic policy gradient algorithm and applications," vol. 26, no. 5, pp. 26–28, 2024. https://doi.org/10.9790/0661-2605032628

[79] N. Darapaneni *et al.*, "Autonomous car driving using deep learning," in *ICSCCC 2021 - International Conference on Secure Cyber Computing and Communications*, May 2021, pp. 29–33. https://doi.org/10.1109/ICSCCC51823.2021.9478090

[80] X. Shi, Y. D. Wong, C. Chai, and M. Z. F. Li, "An Automated machine learning (AutoML) method of risk prediction for decision-making of autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7145–7154, 2021. https://doi.org/10.1109/TITS.2020.3002419

[81] P. Liu and W. Fan, "Extreme gradient boosting (XGBoost) model for vehicle trajectory prediction in connected and autonomous vehicle environment," *Promet - Traffic and Transportation*, vol. 33, no. 5, pp. 767–774, 2021. https://doi.org/10.7307/PTT.V33I5.3779

[82] W. Rong, J. Zheng, Y. Chen, Y. Liu, and Z. Zhang, "Autonomous collision avoidance decision-making method with human-like attention distribution for MASSs based on GMA-TD3 algorithm," *Ocean Engineering*, vol. 330, p. 121118, 2025. https://doi.org/10.1016/j.oceaneng.2025.121118

[83] M. Morari and J. H. Lee, "Model predictive control: Past, present and future," *Computers and Chemical Engineering*, vol. 23, no. 4–5, pp. 667–682, 1999. https://doi.org/10.1016/S0098-1354(98)00301-9

[84] F. Mohseni, S. Voronov, and E. Frisk, "Deep learning model predictive control for autonomous driving in unknown environments," *IFAC-PapersOnLine*, vol. 51, no. 22, pp. 447–452, 2018. https://doi.org/10.1016/J.IFACOL.2018.11.593

[85] M. Bouton, A. Nakhaei, K. Fujimura, and M. J. Kochenderfer, "Scalable decision making with sensor occlusions for autonomous driving," in *Proceedings - IEEE International Conference on Robotics and Automation*, Sep. 2018, pp. 2076–2081. https://doi.org/10.1109/ICRA.2018.8460914

[86] H. Kurniawati, D. Hsu, and W. S. Lee, "SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces," *Robotics: Science and Systems*, vol. 4, pp. 65–72, 2009. https://doi.org/10.7551/mitpress/8344.003.0013

[87] D. Hsu, W. S. Lee, and N. Rong, "A point-based POMDP planner for target tracking," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2644–2650, 2008. https://doi.org/10.1109/ROBOT.2008.4543611

[88] N. Saldi, S. Sanjari, and S. Yuksel, "Quantum Markov decision processes: general theory, approximations, and classes of policies," pp. 1–30, 2024. [Online]. Available: http://arxiv.org/abs/2402.14649

[89] M. S. Ying, Y. Feng, and S. G. Ying, "Optimal policies for quantum markov decision processes," *International Journal of Automation and Computing*, vol. 18, no. 3, pp. 410–421, 2021. https://doi.org/10.1007/s11633-021-1278-z

[90] D. S. Gonzalez, M. Garzon, J. S. Dibangoye, and C. Laugier, "Human-like decision-making for automated driving in highways," in *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, Oct. 2019, pp. 2087–2094. https://doi.org/10.1109/ITSC.2019.8917449.

[91] I. Chadès, L. V. Pascal, S. Nicol, C. S. Fletcher, and J. Ferrer-Mestres, "A primer on partially observable Markov decision processes (POMDPs)," *Methods in Ecology and Evolution*, vol. 12, no. 11, pp. 2058–2072, 2021. https://doi.org/10.1111/2041-210X.13692.

[92] S. James, G. Konidaris, and B. Rosman, "An analysis of Monte Carlo tree search," *31st AAAI Conference on Artificial Intelligence, AAAI 2017*, no. 1983, pp. 3576–3582, 2017. https://doi.org/10.1609/aaai.v31i1.11028.

[93] A. R. Syamil and D. Kum, "Learning homotopy prediction for optimization-based trajectory planners for autonomous driving," *International Journal of Automotive Technology*, vol. 27, no. 1, pp. 59–69, 2025. https://doi.org/10.1007/s12239-025-00257-1.

[94] S. R. Arumugam, P. Sheela Gowr, J. P. Ananth, S. G. Karuppasamy, S. Palani, and J. Elavarasi, "Empowering autonomous vehicles to make challenging options in unexpected circumstances with hybrid learning," *Proceedings of 3rd International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2025*, pp.

1383–1387, 2025. https://doi.org/10.1109/ICSCDS65426.2025.11167771
[95] J. Liang, C. Tan, L. Yan, J. Zhou, G. Yin, and K. Yang, "Interaction-aware trajectory prediction for safe motion planning in autonomous driving: A transformer-transfer learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 10, pp. 17080–17095, 2025. https://doi.org/10.1109/TITS.2025.3588228
[96] R. Chekroun, T. Gilles, M. Toromanoff, S. Hornauer, and F. Moutarde, "MBAPPE: MCTS-built-around prediction for planning explicitly," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 2062–2069, 2024. https://doi.org/10.1109/IV55156.2024.10588457
[97] Y. Liao, G. Yu, P. Chen, B. Zhou, and H. Li, "Integration of decision-making and motion planning for autonomous driving based on double-layer reinforcement learning framework," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 3, pp. 3142–3158, 2023. https://doi.org/10.1109/TVT.2023.3326548
[98] J. Gao, N. Liu, H. Li, Z. Li, C. Xie, and Y. Gou, "Reinforcement learning decision-making for autonomous vehicles based on semantic segmentation," *Applied Sciences (Switzerland)*, vol. 15, no. 3, 2025, doi: 10.3390/app15031323
[99] M. Thoma, "A survey of semantic segmentation," pp. 1–16, 2016. [Online]. Available: http://arxiv.org/abs/1602.06541
[100] S. Mishra and A. Arora, "Double deep Q network with huber reward function for cart-pole balancing problem," *International Journal of Performability Engineering*, vol. 18, no. 9, pp. 644–653, 2022. https://doi.org/10.23940/ijpe.22.09.p5.644653
[101] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," pp. 1–12, 2017, [Online]. Available: http://arxiv.org/abs/1707.06347
[102] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International Journal of Multimedia Information Retrieval*, vol. 7, no. 2, pp. 87–93, 2018. https://doi.org/10.1007/s13735-017-0141-z
[103] S. Chen, M. Wang, W. Song, Y. Yang, and M. Fu, "Multi-agent reinforcement learning-based decision making for twin-vehicles cooperative driving in stochastic dynamic highway environments," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 10, pp. 12615–12627, 2023. https://doi.org/10.1109/TVT.2023.3275582
[104] S. V Albrecht and L. Schäfer, *Multi-agent reinforcement learning: Foundations and modern approaches*. London, England: The MIT Press, 2024.
[105] Z. Zhou, G. Liu, and Y. Tang, "Multi-agent reinforcement learning: methods, applications, visionary prospects, and challenges," vol. 1, no. 1, pp. 1–43, 2023. https://doi.org/10.1109/TIV.2024.3408257
[106] S. Jagan, G. K. Jadhav, C. Priya, P. Adhikary, U. M. Bhatt, and V. Jayasri, "Reinforcement learning-based autonomous collision prediction and avoidance system for vehicles," *Proceedings of the 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, ICSES 2023*, no. December 2023, pp. 1–5, 2023. https://doi.org/10.1109/ICSES60034.2023.10465417
[107] C. Guo, J. Wu, P. Luo et al., "Hybrid reinforcement learning-based collision avoidance algorithm for autonomous vehicle clusters," *IEEE Access*, vol. 13, no. February, pp. 61564–61577, 2025. https://doi.org/10.1109/ACCESS.2025.3553968
[108] A. Ellouze, M. Karray, and M. Ksantini, "A hybrid decision-making framework for autonomous vehicles in urban environments based on multi-agent reinforcement learning with explainable AI," *Vehicles*, vol. 8, no. 1, p. 8, 2026. https://doi.org/10.3390/vehicles8010008
[109] J. Shi, M. Zhang, J. Zhou, and S. Pan, Reinforcement learning decision-making for autonomous driving based on attention mechanism. *InProceedings of the 2024 Guangdong-Hong Kong-Macao Greater Bay Area International Conference on Digital Economy and Artificial Intelligence*, pp. 867-874, 2024. https://doi.org/10.1145/3675417.3675561
[110] M. Wu, F. R. Yu, and P. X. Liu, "A collective and continuous rule learning framework towards enhanced driving safety and decision transparency," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 10, pp. 14275–14290, 2024. https://doi.org/10.1109/TVT.2024.3398661
[111] G. P. Antonio and C. Maria-Dolores, "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. https://doi.org/10.1109/TVT.2022.3169907
[112] N. Kassem, S. Eid, and Y. Elshaer, "Towards fully autonomous driving: classical and deep learning methods novel integration in a hybrid decision-making framework," *IEEE Transactions on Intelligent Vehicles*, vol. 11, no. 4, pp. 477–488, 2026. https://doi.org/10.1109/TIV.2026.3658258
[113] J. Tang, X. Luo, J. Chen, Y. Yuan, and J. Loo, "An ethical decision making algorithm for autonomous vehicles during an inevitable collision," *Proceedings of 2024 4th International Conference on Big Data, Artificial Intelligence and Risk Management, ICBAR 2024*, pp. 1077–1081, 2025. https://doi.org/10.1145/3718751.3718928