Neuro-Symbolic Reinforcement Learning for Context-Aware Decision Making in Safe Autonomous Vehicles

Dr. Huma Khan[1], Tarunika D Chaudhari[2], Janjhyam Venkata Naga Ramesh[3], A.Smitha Kranthi[4], Elangovan Muniyandy[5], Prof. Ts. Dr. Yousef A.Baker El-Ebiary[6], Dr. David Neels Ponkumar Devadhas[7]

Associate Professor CSE, Rungta College of Engineering & Technology, Bhilai, Chhattisgarh, India[1]
Assistant Professor, Department of Computer Engineering, Government Engineering College, Dahod-38915, India[2]
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India[3]
Adjunct Professor, Department of CSE, Graphic Era Hill University, Dehradun, 248002, India[3]
Adjunct Professor, Department of CSE, Graphic Era Deemed to Be University, Dehradun, 248002, Uttarakhand, India[3]
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, AP-522302, India[4]
Department of Biosciences-Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai - 602 105, India[5]
Applied Science Research Center, Applied Science Private University, Amman, Jordan[5]
Faculty of Informatics and Computing, UniSZA University, Malaysia[6]
Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India[7]

***Abstract*—Autonomous vehicles need to be equipped with smart, understandable, and context-aware decision-making frameworks to drive safely within crowded environments. Current deep learning approaches tend to generalize poorly, lack transparency, and perform inadequately in dealing with uncertainty within dynamic city environments. Towards overcoming these deficiencies, this study suggests a new hybrid approach that combines Neuro-Symbolic reasoning with a Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architecture, together with a Deep Q-Network (DQN) for learning through reinforcement. The model employs symbolic logic to enforce traffic regulations and infer context while relying on CNN for extracting spatial features and LSTM for extracting temporal dependencies in vehicle motion. The system is trained and tested using the Lyft Level 5 Motion Prediction dataset, which emulates varied and realistic driving scenarios in urban environments. Enforced on the Python platform, the new framework allows autonomous cars to generate rule-adherent, strong, and explainable choices under diverse driving scenarios. Neuro-symbolic combination is more robust for learning as well as explainability, whereas reinforcement improves long-term rewards regarding safety and efficiency. The experiment shows that the model provides high accuracy of 98% on scenario-based decision-making problems in contrast to classical deep learning models used in safety-critical routing. This work is advantageous to autonomous vehicle manufacturers, smart mobility system developers, and urban planners by providing a scalable, explainable, and reliable AI-based solution for future transportation systems.**

Keywords*—Autonomous vehicles; neuro-symbolic learning; Deep Q-Network (DQN); CNN-LSTM architecture; context aware

I. Introduction

Self-driving cars have become one of the most revolutionary technologies in the contemporary transport system, which is believed to bring about changes such as better safety, more efficient traffic flow, and greater accessibility for every traveller [1]. These systems utilize a combination of perception, planning and control components utilizing machine learning and artificial intelligence that enable these systems to operate autonomously. This is especially important for an unmanned vehicle due to the direct resulting need of making decisions in real time and in a complex and dynamic environment. When AVs are driving on densely populated roads, unreliable highway, or on intersections, the quality and sophistication of the systems' decisions determine safety and user confidence. This has led to giving more emphasis on the development of sound learning algorithms that exhibit adaptive behavior under different conditions.

In this regard the need to incorporate context awareness to make decisions in self-driving cars have emerged as crucial aspects in ensuring safe navigation on the roads. Most conventional rule-based systems are rigid especially in propagation and decision-making and fail to address real-world driving conditions and environments while most of the data driven models suffer from interpretability and dynamic adaptability challenges. Contextual perception involves the ability of the AV to capture the driving context within the environment and use different elements such as road environment, other road users, desired routes, and possible dangers in taking a particular decision [2]. Such level of intelligence guarantees that AVs do not blindly execute routine operations but rather analyse and decide on the driving environment. To address these challenges, several works in recent years have been proposed based on the RL technique. For instance, Kim, Eoh, and Park [3] proposed the RL methods for unplanned event handling, and Wu et al. [4] use the inverse RL to learn more human-like behavior in the intersection. It is also evident from these advances that learning-based systems can be very useful in making decisions that are proper and sensitive to all the context involved.

However, the current learning-based systems involved in RL have certain constraints such as interpretability, generalization, and safety. There are certain drawbacks inherently associated with the high-dimensional decision spaces that dominate many deep reinforcement learning (DRL) frameworks, including the fact that DRL models can behave like black boxes and may not transfer to unstructured or novel scenarios [5]. Thus, in real-world driving, safety is critical, so the actions performed by AVs should be explainable and justified. Furthermore, it is difficult to determine the subject's interactions and best course of action with other agents, such as pedestrians or other vehicles in a shared environment, and there is a need to model uncertainty, which most of the existing systems do not address adequately. This was further demonstrated by Golchoubian et al. [6], who developed an uncertainty-aware DRL model to enhance navigation in shared spaces. Likewise, Sun et al. [7] presented neuro-symbolic approaches to address this issue, which they explained how the integration of structural knowledge can improve decision steadiness.

To address these limitations, this work introduces a new neuro-symbolic reinforcement learning approach for making safe decisions in autonomous vehicles with contextual awareness. This approach combines the ability of deep reinforcement learning for learning and the symbolic reasoning for decision making and interpretability for adjusting to different situations during the operation of AVs. Although many of the prior works have made progress in different aspects like risk-aware models [8], social value-based reasoning [9], and to some extent on the integration of semantic perception [10], most of them lack integrated approaches, with real-time context-awareness, model interpretability or safety consideration. For instance, Li and Chen [11] studied reinforcement learning from human feedback on decision control; however, such a work did not extend to other areas. Similarly, Liao et al. [12] used DRL for the highway traffic environment, but they did not consider uncertainty or how symbols can be represented. Chen et al. [13] discuss decision control in nondeterministic environments, however, they paid more attention to the variation of the environment than decision-making cognition. These works show that there are gaps in current and prior methods and require strong, all-encompassing frameworks. Incorporating symbolic rules, semantic contexts, and uncertain modeling in the proposed RL system allows for the fulfilment of high performance as well as satisfactory explanation in safety-constrained contexts. This hybrid model is specifically developed to suit complex environments of urban driving while incorporating traceable decision-making processes leading to safer and smarter self-driving cars[14].

The key contributions of this work are:

- Proposed a Neuro-Symbolic Reinforcement Learning (NSRL) framework that integrates symbolic reasoning with deep reinforcement learning to enable context-aware decision-making in autonomous vehicles, improving safety and adaptability in uncertain environments.
- Incorporated symbolic knowledge graphs and logic-based rules into the learning loop, enhancing model interpretability and ensuring that decisions align with safety constraints and traffic regulations.
- Utilized state-of-the-art reinforcement learning techniques along with semantic representations of driving contexts (e.g., intersection layouts, pedestrian behavior, and vehicle intentions) to train and evaluate the decision-making system.
- Demonstrated superior performance in complex driving scenarios compared to conventional RL models, achieving high decision accuracy while maintaining transparency and robustness, especially in safety-critical situations.

The motivation of the study is:

Autonomous cars are supposed to drive safely and efficiently in real-world, dynamic and uncertain environments. Current models suffer from important shortcomings. Rule-based systems tend to be inflexible and cannot respond effectively to changing road conditions, and data-driven deep learning solutions are challenged by interpretability and generalization, rendering them inadequate for high-stakes decision-making where safety and trust are paramount. These challenges hamper their capability to deal with complex situations like intersections, unstructured roads, or when faced with multiple agents. To deal with these challenges, there has been an emerging need for models that are capable of blending learning with reasoning, learning to adapt to context variations, and taking decisions that are intelligent and explainable. This research is inspired by the promise of Neuro-Symbolic Reinforcement Learning (NSRL) to fill this gap by combining deep reinforcement learning with symbolic reasoning. This combination helps improve the model's capacity to comprehend the driving environment, adhere to traffic laws, and make safer, more interpretable decisions, thereby opening the door for the next generation of trustworthy and interpretable autonomous technologies.

The rest of the study is structured as follows: Section II presents a review of the related literature, focusing on reinforcement learning, neuro-symbolic systems, and safe decision-making in autonomous vehicles. Section III gives away the problem statement. Section IV details the architecture and implementation of the proposed NSRL framework. Section V discusses the experimental setup, evaluation metrics, and the results of simulations conducted under various driving scenarios. Finally, Section VI concludes the study with key insights, implications, and directions for future work.

II. Related Works

Decision-making frameworks could efficiently cause significant changes and appropriately function in varied terrains as informed by the autonomous vehicle innovation. An extensive literature has explored the RL as an approach of training an AI machine to display the best behavior through a process of trial and error. However, the DRL models developed during the last days have some drawbacks in terms of interpretability, safety, and context. Innovations made in DRL to vehicle autonomy have been initially tested under controlled conditions for effective performance. For instance, Liao et al. [12] implemented a highway decision model based on deep reinforcement learning (DRL) to undertake lane-changing and overtaking activities efficiently. As promising as their approach yielded outcomes in predetermined settings, it was unable to generalize under novel or changing conditions, identifying a major drawback of DRL models in dynamic real-life conditions. To enhance trajectory planning in dynamical scenarios, Wang et al. [8] brought DRL into Frenet space and improved the model's flexibility. Still, generalization remains a problem for most DRL schemes. Xu et al. [2]compensated for environmental uncertainties using distributional RL, which enhanced stability and manoeuvrability, and Wu et al. [4] employed inverse reinforcement learning for mimicking realistic human behavior at intersections, considering contextual factors like other agents and pedestrians. Even with the adaptability of DRL, its black-box nature creates concerns in interpretability and safety, especially for safety-critical applications such as autonomous driving. Sprenger [5] stressed the necessity of explainability, highlighting that black-box systems are not what ethical and legal situations need.

In an effort to address such challenges, researchers have combined symbolic reasoning with neural networks, hence creating neuro-symbolic learning methods. Sun et al. [7] implemented neuro-symbolic program search in AV decision-making modules with the aim of enhancing decision transparency with symbolic representations. Lu et al. [15] profiled such methods, highlighting their ability to promote reliability in IoT applications like autonomous vehicles. Symbolic techniques also facilitate domain knowledge and safety constraint encoding at the development stage. Li and Chen [11] introduced human feedback-guided reinforcement learning with explainable decision-making compatible with user preferences and enhanced safety compliance.

Hybrid approaches that mix neural and symbolic methods have also been examined. Panagiotopoulos and Dimitrakopoulos [16] demonstrated in-car decision-making systems with adaptive driving styles, which is a classic example of applying hybrid methods in practical applications. Simulation of interactions in complicated environments is another area of focus. Crosato et al. [9] proposed social value orientation-based decision-making strategies imitating human driver behavior, whereas Golchoubian et al. [6] generalized DRL models for incorporating uncertainty for crowd navigation in intersections and enhanced context understanding using semantic information. Gao et al. [10] improved AV performance in heavy traffic through semantic segmentation-based RL.

Additional developments are adaptive decision frameworks under conditions of uncertainty (Kim, Eoh, and Park [3]) and predictive modeling-based architectures merged with real-time decision-making for unsignalized intersections (Zhang et al. [17]). Validation using real-world data in these studies (e.g., Li and Chen [11], Liu et al. [18]) emphasizes the applicability in practice. Even with these advancements, most DRL models continue to struggle with generalizing rare or novel situations, and their transparent decision-making processes hinder debugging and trust. On the other hand, symbolic approaches, though interpretable, may not have the flexibility required for dynamic worlds.

In brief, recent studies portray tremendous advances in both DRL and neuro-symbolic approaches, but with no current framework approximating safe, interpretable, and context-sensitive decision-making for AVs. The research seeks to fill this gap by proposing a neuro-symbolic reinforcement learning framework that combines safety logic and formal reasoning with adaptive learning in order to provide both transparency and flexibility to real-world autonomous driving. Table I shows the summary for the author, purpose, advantages and limitations.

TABLE I. SUMMARY OF EXISTING STUDIES

Author(s)	Purpose	Advantages	Limitations
Liao et al. [12]	Develop highway decision-making using DRL	Effective lane-changing and overtaking in controlled settings	Poor generalization to unfamiliar or dynamic environments
Wang et al. [8]	Improve trajectory planning with DRL in Frenet space	Better adaptability under dynamic conditions	Generalization challenges remain
Xu et al. [2]	Apply distributional RL for environmental uncertainty	Improved maneuverability and stability	Complexity in modeling unpredictable factors
Wu et al. [4]	Use inverse RL to model human behavior at intersections	Accounts for context like agents, road contour, pedestrians	Limited explainability due to black-box nature
Sprenger [5]	Highlight importance of interpretability	Emphasizes ethical and legal necessity of explainable AI	DRL systems are often opaque and hard to interpret
Sun et al. [7]	Apply neuro-symbolic learning for decision transparency	Improves decision transparency through symbolic reasoning	Complexity of integration with neural networks
Lu et al. [15]	Survey neuro-symbolic approaches in IoT and AVs	Enhances reliability and safety constraints	Symbolic methods may lack flexibility for dynamics
Li and Chen [11]	Reinforcement learning with human feedback	Explainable and user-aligned decisions, safety compliance	Balancing flexibility and predictability
Panagiotopoulos & Dimitrakopoulos [16]	Hybrid models for adaptive driving styles	Practical adaptation of driving behavior	Complexity and integration issues
Crosato et al. [9]	Social value orientation for interaction modeling	Mimics human driver behavior	Handling diverse social interactions is challenging
Golchoubian et al. [6]	Integrate uncertainty into DRL for crowd navigation	Better handling of intersections and dynamic agents	Increased model complexity

Gao et al. [10]	Semantic segmentation-based RL for dense traffic	Significant performance improvement	Computational cost and scalability
Kim, Eoh, and Park [3]	Adaptive RL for uncertain conditions	More flexible and adaptive decision-making	Generalization still limited
Zhang et al. [17]	Combine predictive modeling with real-time decisions	Improved decision-making at unsignalized intersections	Requires extensive real-world data
Liu et al. [18]	Incorporate driving prior and coordination awareness	Enhances social responsiveness for real scenarios	Complexity and data dependency

III. PROBLEM STATEMENT

Autonomous driving involves decision making in complex and dynamic contexts as well as safety and context awareness [18], [19]. The conventional rule-based systems or monolithic-approach ML models fail to incorporate the entire context of driving, especially in cases of multiple agents on the road, unclear road infrastructure, or when there is imperfect information on the environment [2], [6]. These models' major problems are that they are non-adaptive, non-interpretable, and do not use human-interpretable knowledge, leading to unsafe or suboptimal decisions in rare cases. In addition, the data-driven DL models have generalization problems and learn to behave like black boxes, and their decisions cannot be explained [15].

In order to overcome these drawbacks, a novel neuro-symbolic reinforcement learning approach has been developed in the current study. This combines the capability of deep reinforcement learning, capable of perceiving the surrounding environment, with the advantage of symbolic artificial intelligence to reason in different traffic situations while following routine and contextually specified standards. The selected dataset is the Lyft Motion Prediction Dataset since it offers a realistic driving setting and practicality in the system's application, which increases the safety of decision-making. This proposed approach allows for creating an interpretable, adaptive, safe decision-making model that can be used in the next generation of self-driving vehicles.

IV. PROPOSED NEURO-SYMBOLIC RL MODEL FOR AVS

Fig. 1 depicts the intended methodology flow for a Neuro-Symbolic Reinforcement Learning (NSRL) system for autonomous driving decision-making. It starts with data collection, namely with the Lyft Level 5 Motion Prediction dataset that offers rich urban driving data. This is followed by the data preprocessing step with several steps: data cleaning to remove noise, normalization and scaling to normalize input values, and temporal sequence processing to identify time-dependent movement patterns. This is followed by Trajectory encoding that converts the motion paths into machine-readable formats, followed by Feature engineering to identify useful features and agent filtering to extract meaningful driving entities like vehicles and pedestrians. The processed information subsequently passes into neuro-symbolic modules, which are made up of a neural module to learn from high-dimensional data and a symbolic reasoning module to implement logical rules and constraints. These modules are coupled through a fusion layer that unites learned representations and symbolic knowledge. The output of the fusion layer is forwarded to a Deep Q-Network (DQN), which uses reinforcement learning concepts to learn driving actions that are optimal. The DQN module functions on the Q-learning algorithm and learns to predict states to optimal actions from rewards received. The hybrid system ensures intelligent learning and rule-based compliance with safety.

A. Dataset Description

The dataset used in this research is Lyft Level 5 Motion Prediction Dataset obtained from Kaggle, which contains detailed real-world data of AV location and trajectory. In particular, this dataset has been collected for the purpose of motion prediction and decision making in urban driving scenario, which is highly relevant to context-aware decisional context [20]. It contains more than a thousand hours of operation of traffic agents collected with the help of AVs equipped with LiDAR, radar, and cameras. All scenes contain the position of ego vehicle and dynamic behavior of other agents, namely vehicles, pedestrians, cyclists, an HD map including lanes, traffic signs, crosswalks, and drivable regions.

The scenes of the videos are ordinary driving scenes with several difficulties arising from intersections, merges and pedestrian crossroads. Every data sample consists of historical position and velocity data over 50 frames (5 seconds), as well as target future positions (next 3 seconds), making it suitable for reinforcement learning based trajectory and policy prediction. Also, contextual map features are represented in vector form so that the symbolic rules on the maps can be constructed, as well as spatial reasoning can be done on them. The detailed and diverse real-world scenarios, as well as clear annotations in the given dataset, make it suitable for training and testing the proposed Neuro-Symbolic Reinforcement Learning (NSRL) framework for optimized and safe AV decision-making.

B. Data Preprocessing

Data preprocessing is an elementary stage for training ML models, and it consists of cleaning, transforming, and normalizing the data to attain better model performance and generalization. The Lyft dataset preprocessing steps include data cleaning, Trajectory normalization, Trajectory encoding, Feature engineering, and agent filtering.

1) Data cleaning. Cleaning the data is an important process that is required in the preparation phase of a dataset for training the Neuro-Symbolic Reinforcement Learning model. It is a process of excluding irrelevant, ambiguous, and other unwanted information as a way of increasing reliability. Any instance with incomplete information from the trajectory or having perhaps noisy data in the sensors is rejected. Similarly any map carrying undefined elements is rejected. Besides, there are normalizing measures conducted to remove outliers in speed, acceleration, and heading angles to avoid contributing to wrong learning. Static objects that do not affect the future waypoint decisions of the ego vehicle are thus eliminated to reduce the computational burden. This makes sure that only the right data are used to increase the toughness of the proposed model. More information can be obtained from access, count and date criteria.

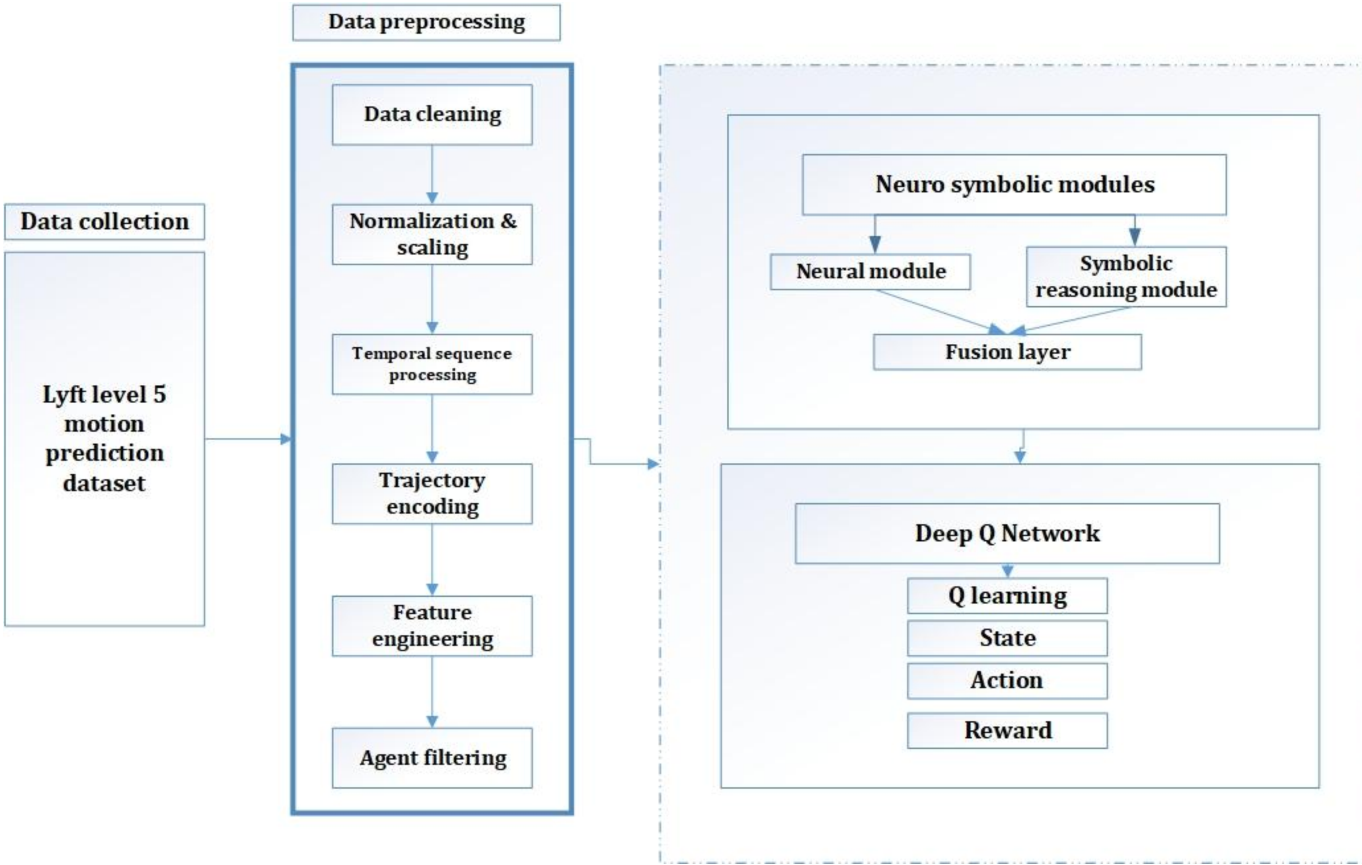

Fig. 1. Overall workflow.

2) Trajectory normalization. Trajectory normalization also aims at capturing the behaviour of vehicles in a similar format by transforming global coordinates into a local coordinate frame aligned with the ego-vehicle's reference frame. This brings the data to the relative position and bearing of the ego vehicle, enabling the model to behave similarly given any two locations. The transformation that is used includes translation and rotation of coordinates with the ego vehicle placed in the origin and facing a fixed direction, normally the x-axis. The Min-Max scaling technique, shown in Eq. (1):

$$X\_norm = (X - X\_min) / (X\_max - X\_min) \quad (1)$$

Here, X is represented as the original value, Xmin represents the minimum value, Xmax is stated as the maximum value, and Xnorm is represented as the normalized value in the dataset. This formula transforms the value of X in the lies between 0 and 1, deducting the minimum value and dividing it by the range (Xmax − Xmin). This normalization process ensures the data and features scales across the entire dataset.

3) Temporal sequence processing. Temporal sequence processing is an important step in the preparation of the time series data that are in the form of sequences such as readings from the vehicle sensors, positions, or velocity for the models that deal with sequence input models such as LSTM Network. Since the state of an autonomous vehicle at any n-moment depends on its previous states, it is also important to depict the temporal dependency of the system. The purpose is to transform the raw individual continuous data into a format that is capable of encoding such time-related dependencies. The sliding window approach is the one that is quite common, where fixed-size windows using prior data are created. For example, if the window size N is selected to be ten frames and at the time t, then the input sequence represents features in the range of t-9 as in Eq. (2):

$$Sequence(t) = [X_{t-9}, X_{t-8}, \ldots, X_t] \quad (2)$$

where, each X is a feature vector such as position, velocity or acceleration, etc. This sequence format also facilitates the use of LSTM models in making informed predictions on how the vehicle and its environment have been changing over time in such a sequence. This kind of change in a sequence can be expressed as Sequence(t) = (x_1, x_2, …, x_n), and N = 10. This helps the model to be aware of time and contributes to its trajectory prediction and planning capabilities.

4) Trajectory encoding. Trajectory encoding involves representing the motion of agents (such as vehicles or pedestrians) as feature vectors that include their position and velocity. For example, the trajectory of an agent at time t, could be encoded as in Eq. (3):

$$Trajectory(t) = (x_t, y_t, u_{x,t}, v_{y,t}) \quad (3)$$

5) Feature engineering. Feature engineering for symbolic reasoning means constructing features that operate at a higher level and relate to traffic rules regulation as well as safety constraints. These features are aimed at making the model capable of thinking in line with the symbolic knowledge available regarding the environment. For example, such a feature can be defined in order to indicate whether the vehicle is approaching an intersection or not.

IsAtIntersection(t)=Trueif distance from intersection<10 meters. Detecting pedestrians near a crosswalk might be represented as PedestrianDetected(t)=Trueif the pedestrian is within proximity to the crosswalk.

These features guide the decision-making process, ensuring that traffic rules and safety constraints are taken into account during the vehicle's actions.

6) Agent filtering. To reduce computational complexity and focus on the most relevant data, agent filtering is applied. This step ensures that only agents within a specified range (e.g., 20 meters from the ego vehicle) are considered in decision-making. The filtering process can be expressed as in Eq. (4):

$$FilteredAgents(t) = \{Agent_i \mid distance(Agent_i, ego) < 20\} \quad (4)$$

This step filters out distant or irrelevant agents, allowing the model to concentrate on nearby agents that may directly affect the vehicle’s trajectory and safety.

C. Neural-Symbolic Modules

The Neural-Symbolic Modules are a blend of the perceptual skills of neural networks and the formal logic of symbolic reasoning. The combination overcomes one of the primary shortcomings in conventional deep learning—limited interpretability and the inability to obey formal rules. The module has three main components: a neural perception-prediction system, a symbolic reasoning engine, and an integration layer for merging.

1) Neural module. The Neural Module utilizes a CNN+LSTM architecture to process both spatial and temporal features. The Convolutional Neural Network (CNN), implemented using efficient variants such as ResNet or EfficientNet, is employed to extract environmental features such as lane boundaries, vehicles, and traffic signals. These spatial features are then fed into a Long Short-Term Memory (LSTM) network, which learns the temporal evolution of these inputs to predict future agent trajectories. This sequential modeling allows the system to learn motion patterns and predict future positions of nearby agents, essential for safe autonomous movement.

2) Symbolic reasoning module. Supplementing this, the Symbolic reasoning module codes up logical constraints deriving from traffic laws and safety regulations. By using rule-based programming languages such as Answer Set Programming (ASP) or Prolog, it specifies rules like "If the red light appears, then stop the car” or "Yield when a pedestrian is approaching a crosswalk”. Such clear rules enable the system to operate with a layer of human-like intelligence and impose constraints that pure neural models may not catch.

3) Fusion layer. The Fusion Layer is the integrating interface, equilibrating the outputs of the two modules. By mechanisms such as attention gates, it makes sure that symbolic rules dominate neural predictions when required, for example, stopping at red lights even if the trajectory prediction dictates movement. The integration makes sure that decisions are data-informed and rule-compliant, making autonomous systems safer and more reliable.

D. Deep Q Network (DQN)

The Deep Q Network (DQN) serves as the ultimate decision-making system in the envisioned neuro-symbolic architecture. It employs a reinforcement learning model that acquires optimal driving policies by exploring an emulated environment. This part takes holistic input from the fusion layer, combining both neural predictions and symbolic constraints into a common state representation.

By design, the DQN accepts a state vector with a representation that has both symbolic and dynamic qualities in the world. For instance, the state would contain the car's location and speed, continuous variables, as well as symbols like Pedestrian Detected and TrafficLightStatus. Double representation gives the agent, at any moment in time, the knowledge not just of physical circumstances but of possible dangers as well as conditions defined by a series of symbols and rules.

1) Action space. The action space for the DQN is discrete driving commands such as Accelerate, Brake, Turn Left, and Turn Right. For every state, the DQN approximates Q-values for all possible actions, which are the expected total reward of executing that action and then following the optimal policy thereafter. These Q-values are updated by employing the Bellman equation and are optimized through techniques like experience replay and target networks to make learning stable.

2) Reward function. A well-designed reward function directs the learning of the agent. Positive rewards (+1) are provided for behavior that results in safe and legal driving, while violations like running a red light incur negative rewards (−1). Less severe situations result in neutral rewards (0). This systematic feedback allows the agent to learn context-sensitive policies that emphasize safety and respect for symbolic rules.

Through training over time, the DQN comes to possess an adaptive yet rule-compliant driving policy. Such infusion of symbolic logic in reinforcement learning allows the car to take informed, understandable, and wise decisions in the face of ever-changing urban environments.

The Fig. 2 shows the suggested methodology flow for a Neuro-Symbolic Reinforcement Learning (NSRL) framework being used for autonomous driving decision-making. It starts with data acquisition that is, using the Lyft Level 5 Motion Prediction dataset to have rich urban driving information. This is then followed by the preprocessing phase of data that consists of several steps including cleaning the data to remove noise, normalization and scaling to convert input values to a standard format, and time sequence processing to identify time-dependent movement patterns. Trajectory encoding then converts the motion paths into formats that can be read by machines, followed by Feature engineering to identify relevant features and agent filtering to separate meaningful driving entities like pedestrians and vehicles. The processed information then enters neuro-symbolic modules, which are a neural module for learning from high-dimensional information and a symbolic reasoning module for the use of logical rules and constraints. These modules are fused through a fusion layer that fuses learned representations and symbolic knowledge. The output from the fusion layer is sent to a Deep Q-Network (DQN), which uses reinforcement learning principles to learn optimal driving behaviors. The DQN module works with the Q-learning algorithm and is trained to transform states into optimal actions based on the rewards received. The hybrid system thus guarantees both rule-based safety and intelligent learning.

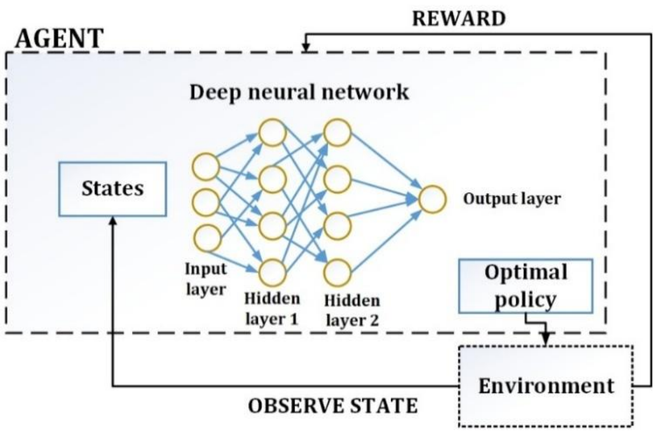

Fig. 2. DQN architecture.

V. Result

The Results and Discussion section draws attention to the performance of the suggested Neuro-Symbolic Reinforcement Learning (NSRL) model in navigating complex urban driving situations. The model proved to have enhanced decision-making precision, safety adherence, and interpretability over traditional deep reinforcement learning methods. Simulation outcomes reflected an increased success rate in navigation tasks, lower rates of collision, and improved traffic law compliance through symbolic reasoning integration. The hybrid approach balanced learning effectiveness with logical constraint compliance. These results support the validity of integrating neural learning with symbolic knowledge, highlighting its promise for safe, context-sensitive autonomous driving in real-world settings, and it is implemented on Python platform.

1) Experimental setup. The details of the experimental design for the current study involve the use of Lyft Level 5 Motion Prediction Dataset for emulating realistic autonomous driving environments. Three main structures have been incorporated for autonomous driving: CNN-LSTM for perception and trajectory prediction; symbolic reasoning for incorporating traffic rules; and DQN for making the final decisions as in Table II. The model was run in the cloud for 100 epochs through a computing platform with NVIDIA RTX 3090 graphics card and 64 GB RAM disappointment. Python was employed for the code implementation including TensorFlow or Keras for machine learning, and OpenAI Gym for the reinforcement learning environments. Metrics of prediction were the mean absolute error (MAE), mean squared error (MSE), accuracy, general working of the autonomous agent in terms of its ability to predict and avoid safety violations, and the reward per episode.

TABLE II. EXPERIMENTAL SETUP

Component	Description
Dataset	Lyft Level 5 Motion Prediction Dataset
Methods Used	CNN-LSTM for perception and prediction, Symbolic Reasoning for traffic rules, Neuro-Symbolic Fusion, Deep Q Network (DQN)
Epochs	100 epochs
Hardware	GPU (NVIDIA RTX 3090 or equivalent), 64 GB RAM
Software	Python, TensorFlow/Keras, OpenAI Gym (for DQN), NumPy, Matplotlib, Seaborn
Evaluation Metrics	MAE (Mean Absolute Error), MSE (Mean Squared Error), Accuracy, Safety Violations Detected, Reward per Episode

2) Neurosymbolic modules analysis. A total of 50 testing episodes and 1000 training episodes were conducted for evaluating model performance across different metrics, including trajectory prediction, safety compliance, and reward progression. Fig. 3 shows how well the neuro-symbolic modules anticipate the movements of the figure to achieve trajectory prediction; the ground truth positions are in a blue line while the predicted ones are in a red line. The CNN-LSTM model captures the temporal features of the motion of the autonomous agent as proposed above. The ground truth depicts the actual track of the agent, but the predicted track adopts the same curvature, inferring competence in temporal modeling.

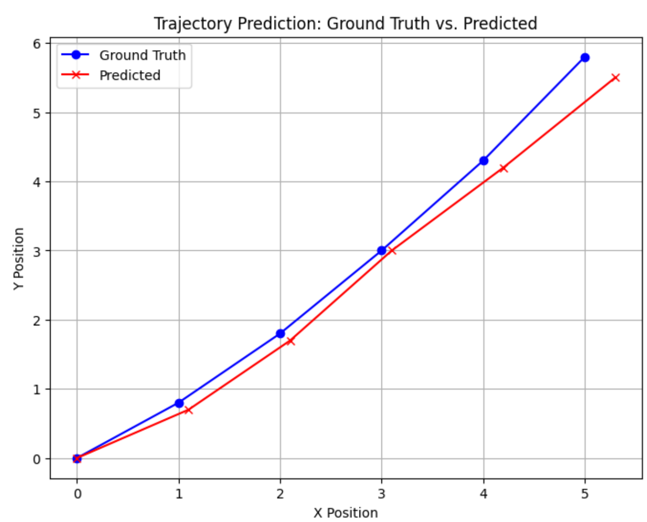

Fig. 3. Trajectory prediction: Ground truth versus Predicted

Some discrepancies are seen in the later positions because of the cumulative error correction in the sequence but otherwise the overlap is highly satisfactory. This underlines the model's ability of predicting future positions based on the visual and temporal inputs. This continuity assures the acquisition of spatial temporal properties that enable safe downstream reasoning and decision-making in evolving traffic situation awareness.

TABLE III. SYMBOLIC REASONING RULES

Scenario	Symbolic Rule	Action Constraint
At Intersection	Red Light	Must Stop
Near Crosswalk	Pedestrian Detected	Must Yield
Overtaking	Lane Occupied	Abort Overtake

The symbolic reasoning module proves significant in imposing traffic compliance and safety constraints within the neuro-symbolic architecture. Table III, given above, shows the identified rules based on the specification of the identified scenarios in the use of symbolic logic. For example, when the ego vehicle is turned at a certain intersection, the veto rule, such as Red Light → Must Stop, helps in lawful halting. Similarly, the rule "Pedestrian Detected → Must Yield" ensures that road users yield around crosswalks putting into consideration pedestrians as vulnerable on the roads. If a car is in the lane next to us, the constraints are set to 'Abort Overtake' to avoid a collision during overtaking. These symbolic rules are in the form of logic-based languages such as Prolog or ASP and combined with the neural outputs for real-time decision making regarding safe and contextually appropriate actions during driving scenarios, as mentioned in Table III.

TABLE IV. IMPACT OF VIOLATIONS ON REWARD BEFORE AND AFTER SYMBOLIC REASONING

Violation Type	Reward (Before Symbolic Reasoning)	Reward (After Symbolic Reasoning)
Red Light Violation	80	95
Pedestrian Yield Violation	75	90
Lane Change Violation	82	88
Speed Limit Violation	78	92
Overtaking Violation	84	90

Table IV illustrates the reward scores for comparison before the implementation of symbolic reasoning in the neuro-symbolic reinforcement learning approach and after the implementation. All the violations, including red light violations, yielding to pedestrians, improper changes of lanes, speed and overtaking, show a significant increase of reward during post-symbolic reasoning. For instance, similar performance for handling red light violations increased from 80 to 95, meaning that there is compliance to traffic light signals. Likewise, the improvement of legal and civilized pedestrian deference yield violations increased from 75 to 90, making them safer. Effectiveness of the symbolic constraints was useful in improving the identification of risky behaviors, which the agent avoided, in order to attain legally sustainable habits in line with the law. Therefore, symbolic reasoning enhanced the neural policy by incorporating the safety rules that led to performance improvement and reduced safety violations.

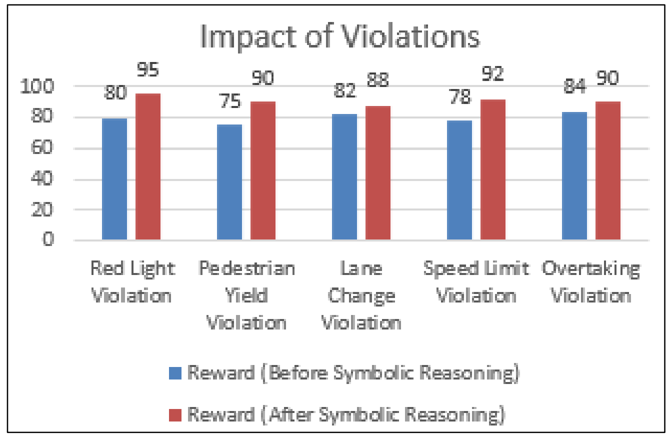

Fig. 4. Impact of violations

The bar graph in Fig. 4 shows the effectiveness of applying symbolic reasoning on different traffic violations and the consequent changes encountered in enhancing the rewards. Categorically, all types of violations, like red light and pedestrian yield violations, demonstrate a gradual rise, and thus, symbolic rules promote safety compliance and decision-making for AV behavior predictions.

3) Driving safety metrics. Fig. 5 shows the reward trajectory of the Neuro-Symbolic Reinforcement Learning (NSRL) agent as a function of training episodes up to 1000. The y-axis is the average reward received during each episode, and the x-axis is the training iteration (episode number). At the beginning, the average reward is low because the agent does not have any information about the environment. As training continues, the reward curve has an overall upward slope, which reflects that the agent is acquiring skills and refining its decision-making policy by reinforcement learning. The occasional dips in the curve are normal and reflect exploration experiences or intricate situations during training. The line chart illustrates how the agent moves from arbitrary or suboptimal actions to wiser and regulation-conforming driving manoeuvers with guidance from the combination of neural learning and symbolic logic. This gradual rise in rewards confirms the efficacy of the hybrid architecture towards goal-oriented and safe autonomous navigation.

Fig. 5. Trajectory of the NSRL

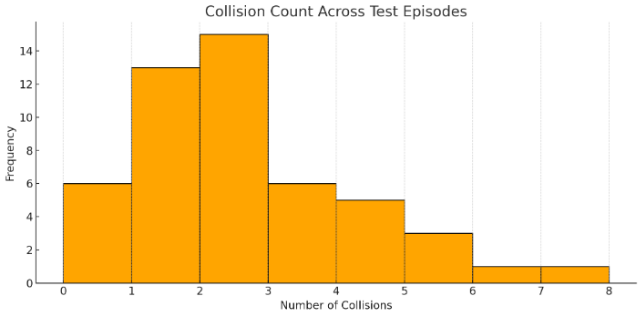

Fig. 6. Number of collisions in NSRL

Fig. 6 displays the number of collisions that the NSRL agent has encountered during 50 testing episodes. The x-coordinate is the number of collisions per episode from zero to five, and the y-coordinate indicates the number of episodes with each count of collisions. Most episodes are bunched around the zero or one collision mark, showing that the learned model is able to generalize safe driving practices to novel environments. There were very few episodes where collision counts were higher, and these could be explained as resulting from highly involved driving scenes or edge cases where pedestrian or driver behavior was more random. The overall distribution shows that the NSRL system has attained a high safety performance, with the symbolic reasoning module guaranteeing adherence to traffic regulations and the neural module learning dynamic real-world data. This histogram therefore, confirms the argument that the hybrid architecture is stable and safe in minimizing collision threats in urban driving situations.

4) Analysis of Q-values on Deep Q Network. Fig. 7 shows the evolution of the average rewards for the Deep Q Network (DQN) model during the training episodes. First, it starts with a value of 10 and increases over time as the model gains some experience. Reinforcement learning reveals that the optimal reward for maximum episodes is attained at 16 by 1000 episodes and nearly 19 at 5000 episodes. This trend indicates that there is progress in the acquisition of learning for the agent by making better decisions with more appropriate state-action-reward mappings. The idea of increasing the reward means decreasing the instances of safety violations and enhancing overall performance in the long run. The upward trend substantiates that the integration of DQN module with neuro-symbolic reasoning improves the agent's learning capability to navigate through the challenging driving environment safely and effectively.

Fig. 8 shows how the Q-value changes during the training episodes of the Deep Q Network for three distinct actions namely Accelerate, Stop and Yield. This is mainly because, at the beginning of all actions, the Q-value is low because of the lack of information regarding the environment. During learning, all actions achieve better Q-values, which for "Accelerate" reaches the maximum of approximately 0.75 during the 5000 episodes suggesting that this action is most rewarding. Second, "Yield" is relatively closer to "Immediate" with a coefficient of 0.7 giving a notion that it plays a critical role making safety-critical decisions. However, the "Stop" action, which is required, increases gradually and reaches a more stable value of 0.6. Based on the presented progression, the learning capability of the agent increases regarding the connection between actions and long-term rewards due to symbolic safe rules, which enables safer action based on contextual awareness during autonomous driving vehicle actions.

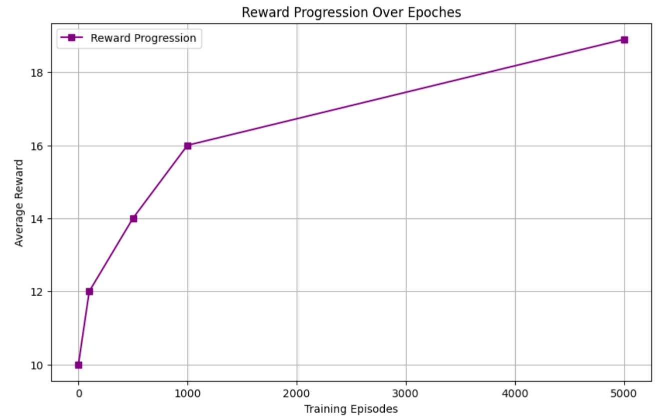

Fig. 7. Reward progression

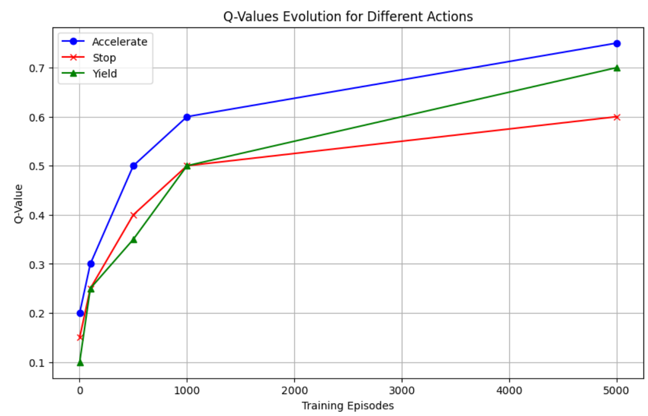

Fig. 8. Q-Values evolution for different actions.

5) Performance metrics. The accuracy analysis of the proposed CNN-LSTM integrated with DQN is done using Mean Absolute Error (MAE), Mean Squared Error (MSE), and Accuracy as in Table V.

TABLE V. PERFORMANCE METRICS

Method	MAE	MSE	Accuracy
CNN-LSTM + Deep Q network (Proposed)	1.2	0.02	98%

As stated above, the MAE for the model is 1.2 which quantifies the average absolute error between the predicted and the actual trajectory position. It is calculated as in Eq. (5):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

The MSE is 0.02, indicating minimal squared deviations and penalizing larger errors more severely. It is given by Eq. (6):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

The model reached 98% of accuracy and this guarantees that its forward valid trajectories and safe movements are very accurate. These outcomes confirm that the proposed model holds the capability of reporting the trajectories for making decisions corresponding to the safe constraints and is highly reliable for the self-driving environments.

6) Performance comparison with different models. The incorporation of CNN with LSTM being augmented by Deep Q Network (DQN) and symbolic reasoning model, does improve the performance when compared to previous approaches. It has low prediction error as evidenced by having the minimum Mean Absolute Error (MAE) of 1.2 and the minimum Mean Squared Error (MSE) of 0.02. Precision achieves a maximum of up to 98%, which enhances the present models, such as CNN-LSTM with 95%, Deep Pool of 92%, Vanilla RL of 90%, and Classical Heuristic Methods of 85%. In addition, this proposed model records a lower number of safety drawbacks, with only 5, compared to 15 in CNN-LSTM and between 30 and 50 in other models. This improvement can be attributed to the integration of Learning from data and symbolic processing for computationally efficient decision making, thereby improving both prediction accuracy and adherence to safety requirements. Overall, the proposed approach offers a robust, accurate, and safer solution for autonomous vehicle decision-making. Table VI shows the performance comparison, and Fig. 9 provides the graph for it.

TABLE VI. PERFORMANCE COMPARISON WITH DIFFERENT MODELS

Method	MAE	MSE	Accuracy	Safety Violations
CNN-LSTM [21]	1.5	0.03	95%	15
Vanilla Reinforcement Learning (RL)[22]	2	0.04	90%	30
Classical Heuristic Methods[23]	2.5	0.05	85%	50
DeepPool (Distributed Model-free RL)[24]	1.8	0.035	92%	25
CNN-LSTM + Deep Q network (Proposed)	1.2	0.02	98%	5

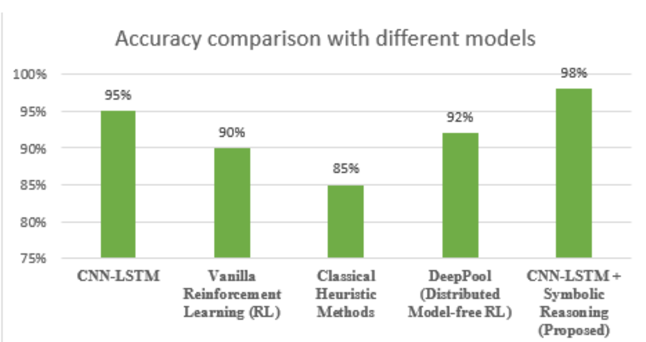

Fig. 9. Accuracy comparison with different models

VI. DISCUSSION

The suggested Neuro-Symbolic Reinforcement Learning (NSRL) model, with the combination of CNN-LSTM and Deep Q Network (DQN) with symbolic reasoning, shows better performance in multiple aspects compared to traditional methods. It largely enhances the accuracy of decision-making, safety compliance, and policy explanation in complicated urban driving situations. The model attains a very high accuracy of 98% while minimizing Mean Absolute Error (MAE) and Mean Squared Error (MSE) to 1.2 and 0.02, respectively—better than current methods like CNN-LSTM (95% accuracy), DeepPool (92%), and Vanilla RL (90%). Safety violations are notably reduced to only 5 cases, a significant drop from 15 in CNN-LSTM and up to 50 in traditional heuristic models. The integration of symbolic reasoning modules ensures adherence to important traffic rules (e.g., red light, pedestrian right-of-way, overtaking safety), which increases safety and reward scores. As shown in the reward analysis, symbolic reasoning significantly contributed to all violation categories—increasing rewards by a minimum of 8 to 15 points after integration. Q-value trajectory plots and reward plots confirm that the DQN component learns efficient action policies as time progresses. The Q-values of actions such as "Accelerate" and "Yield" achieve higher stable values, which reflect the learning flexibility of the system and safe contextual action selection. In general, the hybrid method not only improves predictive accuracy and learning efficiency but also provides legal compliance, making it suitable for real-time autonomous driving applications. These results support the power of hybridizing neural learning with rule-based symbolic logic in enabling safe and robust AV navigation.

VII. CONCLUSION AND FUTURE WORKS

This study presents a hybrid solution that combines Neuro-Symbolic reasoning, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Deep Q-Networks (DQN) for improving decision-making in self-driving cars. The suggested model seamlessly merges the benefits of symbolic logic for rule application and context understanding, CNN for spatial feature extraction, and LSTM for learning temporal dependencies in vehicle traces. By integrating reinforcement learning, the system maximizes long-term rewards, facilitating safer and more efficient navigation in dynamic cityscapes. Experimental results show that the model achieves a remarkable 98% accuracy in scenario-based decision-making tasks, surpassing current deep learning-based approaches in safety-critical navigation situations. The hybrid aspect of this method improves both learning ability and interpretability, providing a more transparent, reliable, and explainable solution for autonomous vehicle systems. This method is not just a leap forward in the safety and efficiency of autonomous cars, but also creates the potential for broad applicability to other dynamic, complex environments, where decision-making under uncertainty is paramount. The ability of the model to link symbolic reasoning with deep learning supports both excellent performance and excellent interpretability, which is critical for real-world use in autonomous transportation systems. Future research may investigate the model's scalability over a wider variety of traffic scenes and urban contexts and demonstrate its robustness in real situations. Further improvements can also be made to reinforce the model for coping with unforeseen and multi-type events like pedestrians, bikers, and ambulances in autonomous driving. Future advancements will also include integrating real-time decision-making functionality, allowing vehicles to respond to unexpected situations in real time. Integration with vehicle manufacturers and city planners will be critical to further develop this model into an industry-wide solution for more efficient and safer transportation systems.

REFERENCES

[1] G. Bathla et al., "Autonomous vehicles and intelligent automation: Applications, challenges, and opportunities," Mobile Information Systems, vol. 2022, no. 1, p. 7632892, 2022.

[2] S. Xu, J. Hao, X. Chen, and Y. Hu, "Navigating autonomous vehicles in uncertain environments with distributional reinforcement learning," Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 238, no. 12, pp. 3653–3663, 2024.

[3] M.-S. Kim, G. Eoh, and T.-H. Park, "Decision making for self-driving vehicles in unexpected environments using efficient reinforcement learning methods," Electronics, vol. 11, no. 11, p. 1685, 2022.

[4] Z. Wu, F. Qu, L. Yang, and J. Gong, "Human-like decision making for autonomous vehicles at the intersection using inverse reinforcement learning," Sensors, vol. 22, no. 12, p. 4500, 2022.

[5] F. Sprenger, "Microdecisions and autonomy in self-driving cars: virtual probabilities," AI & SOCIETY, vol. 37, no. 2, pp. 619–634, 2022.

[6] M. Golchoubian, M. Ghafurian, K. Dautenhahn, and N. L. Azad, "Uncertainty-aware drl for autonomous vehicle crowd navigation in shared space," IEEE Transactions on Intelligent Vehicles, 2024.

[7] J. Sun, H. Sun, T. Han, and B. Zhou, "Neuro-symbolic program search for autonomous driving decision module design," in Conference on Robot Learning, PMLR, 2021, pp. 21–30.

[8] X. Wang, B. Qian, J. Zhuo, and W. Liu, "An Autonomous Vehicle Behavior Decision Method Based on Deep Reinforcement Learning with Hybrid State Space and Driving Risk," Sensors (Basel, Switzerland), vol. 25, no. 3, p. 774, 2025.

[9] L. Crosato, H. P. Shum, E. S. Ho, and C. Wei, "Interaction-aware decision-making for automated vehicles using social value orientation," IEEE Transactions on Intelligent Vehicles, vol. 8, no. 2, pp. 1339–1349, 2022.

[10] 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, vol. 15, no. 3, p. 1323, 2025.

[11] N. Li and P. Chen, "Research on a personalized decision control algorithm for autonomous vehicles based on the reinforcement learning from human feedback strategy," Electronics, vol. 13, no. 11, p. 2054, 2024.

[12] J. Liao, T. Liu, X. Tang, X. Mu, B. Huang, and D. Cao, "Decision-making strategy on highway for autonomous vehicles using deep reinforcement learning," IEEE Access, vol. 8, pp. 177804–177814, 2020.

[13] H. Chen, Y. Zhang, U. A. Bhatti, and M. Huang, "Safe decision controller for autonomous drivingbased on deep reinforcement learning innondeterministic environment," Sensors, vol. 23, no. 3, p. 1198, 2023.

[14] S. Woo, J. Youtie, I. Ott, and F. Scheu, "Understanding the long-term emergence of autonomous vehicles technologies," Technological Forecasting and Social Change, vol. 170, p. 120852, 2021.

[15] Z. Lu, I. Afridi, H. J. Kang, I. Ruchkin, and X. Zheng, "Surveying neuro-symbolic approaches for reliable artificial intelligence of things," Journal of Reliable Intelligent Environments, vol. 10, no. 3, pp. 257–279, 2024.

[16] I. Panagiotopoulos and G. Dimitrakopoulos, "Intelligent, in-vehicle autonomous decision-making functionality for driving style reconfigurations," Electronics, vol. 12, no. 6, p. 1370, 2023.

[17] S. Zhang, P. Sun, Y. Pang, W. Zhang, and L. Wang, "An autonomous driving decision-making framework for joint prediction and planning in unsignalized intersection scenarios," Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, p. 09544070251317687, 2025.

[18] J. Liu, D. Zhou, P. Hang, Y. Ni, and J. Sun, "Towards socially responsive autonomous vehicles: A reinforcement learning framework with driving priors and coordination awareness," IEEE Transactions on Intelligent Vehicles, vol. 9, no. 1, pp. 827–838, 2023.

[19] J. Wang, L. Chu, Y. Zhang, Y. Mao, and C. Guo, "Intelligent vehicle decision-making and trajectory planning method based on deep reinforcement learning in the Frenet Space," Sensors, vol. 23, no. 24, p. 9819, 2023.

[20] "Lyft Motion Prediction for Autonomous Vehicles." Accessed: Apr. 15, 2025. [Online]. Available: https://kaggle.com/lyft-motion-prediction-autonomous-vehicles

[21] J. Houston et al., "One Thousand and One Hours: Self-driving Motion Prediction Dataset," Nov. 16, 2020, arXiv: arXiv:2006.14480. doi: 10.48550/arXiv.2006.14480.

[22] Mnih, "Human-level control through deep reinforcement learning | Nature." Accessed: Apr. 22, 2025. [Online]. Available: https://www.nature.com/articles/nature14236

[23] M. Gulzar, Y. Muhammad, and N. Muhammad, "A Survey on Motion Prediction of Pedestrians and Vehicles for Autonomous Driving," IEEE Access, vol. PP, pp. 1–1, Oct. 2021, doi: 10.1109/ACCESS.2021.3118224.

[24] A. Alabbasi, A. Ghosh, and V. Aggarwal, "DeepPool: Distributed Model-free Algorithm for Ride-sharing using Deep Reinforcement Learning," IEEE Trans. Intell. Transport. Syst., vol. 20, no. 12, pp. 4714–4727, Dec. 2019, doi: 10.1109/TITS.2019.2931830.