2026 2nd International Conference on Intelligent Systems and Computational Networks (ICISCN)

Intelligent Traffic Signal Optimization using Multi-Agent G - DQN with Adaptive Reward Optimization

1st Mogili Venu
Department of CSE (Data Science)
Malla Reddy Engineering College for Women (Autonomous)
Hyderabad, India
mogilivenu556@gmail.com

2nd Manjula Sanjay Koti
Department of Computer Science Engineering (Data Science)
Dayananda Sagar Academy of Technology and Management
Bengaluru, India
manjula@dsatm.edu.in

3rd R. Deepa
Department of Computer Science and Engineering
Vels Institute of Science, Technology and Advanced Studies
Chennai, India
deepar.se@vistas.ac.in

4th K. Jayakumar
Department of Electrical and Electronics Engineering
J.J. College of Engineering and Technology
Trichy, India
rkjkumar70@gmail.com

5th A. Abhinav Reddy
Department of Computer Science and Engineering (AI&ML)
Sreyas Institute of Engineering and Technology
Hyderabad, India
reddyabhinav466@gmail.com

2026 2nd International Conference on Intelligent Systems and Computational Networks (ICISCN) | 979-8-3315-8871-7/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICISCN67954.2026.11566509

Abstract*—Background: in recent years, the Deep Reinforcement Learning (DRL) advanced traffic signal optimization framework is to address the difficulties of inefficient signal control and urban traffic congestion. Challenges: however, existing traffic management system include actuated control methods and fixed-time, which lacks the adaptability and complex traffic conditions, leads to environmental impact, fuel consumption and increased delay. Proposed Methodology: to overcome these challenges, the proposed Multi-Agent Grouping Deep Q Network (GDQN) with Adaptive Reward Optimization (ARO) used the multi agent system where every intersection is demonstrated as the individual and independent agent that learns ideal control policies through environment interaction. Moreover, real time traffic parameters such as traffic density, waiting time, vehicle count and queue length used as input states. This framework uses the GDQN to decrease the state space complexity through the classification of traffic signals into discrete levels, in enhancing the convergence speed and learning efficiency. Further, an adaptive reward optimization adjusts dynamically based on traffic conditions, allows more context aware and responsive decision making. Furthermore, multi-agent coordination improves the global traffic and scalability through the limited information across neighbouring intersections. Results: Experimental results represent that proposed GDQN with adaptive reward mechanism outperformed the existing YOLOv5 method in terms of precision of 99.5%, recall of 98.6% and f1-score of 98.8%. Therefore, this approach is effective and scalable approach for real-time traffic management in smart city environments.

Keywords*—*adaptive reward optimization, intelligent traffic signal optimization, multi-agent coordination, multi-agent grouping deep q network, traffic management system

I. INTRODUCTION

In recent years, rapid development and the sustained expansion in vehicle ownership have increasing world-widely due to the heavy traffic congestion in smart cities. Moreover, for city planners effective traffic maintenance has become a critical challenge leading to maximization of travel time and also results in economic loss., environmental pollution, fuel wastage [1-5]. Correspondingly, for controlling traffic flow traffic signal control plays a major role at intersections where improper signal coordination significantly affects the overall transportation efficiency. The intelligent traffic systems are progressively being discovered with the enhancement of Artificial Intelligence (AI) to advance the adaptability and real-time decision-making in flexible different traffic conditions [6-8]. Further, traditional traffic signal management techniques such as actuated control systems and fixed-time control have possessed limited capability to manage time-varying and complex traffic patterns. Subsequently, time-critical control systems heavily rely on established schedules that fail to adapt the real-time traffic situations making them unproductive during peak travel hours and unpredictable flow. Consequently, powered mechanisms often lack global optimization while it modifies the behavior based on traffic demand and also fail to deliberate max-term traffic dynamics. Similarly, these methods naturally rely on simple mathematical assumptions and models which limit their performance and scalability in wide-range heterogeneous urban networks [9-10]. Correspondingly, the control mechanisms were necessary because of increasing complexity of urban traffic systems.

Moreover, an effective solution was offered by the Deep Reinforcement Learning (DRL) by allowing AI agents to learn most effective strategies through interaction with the environment. Further, this research aims to employ DRL to design effective data-driven and self-adaptive framework for minimizing congestion and enhancing both commuter experience and traffic efficiency [11-12]. This method holds substantial practical value within the realm of smart city development where smart transportation systems are important for sustainable urban development. To estimate the ideal action-value function the DQN-based methods usage Deep Neural Networks (DNNs) for allowing flexible signal control decisions. Moreover, these approaches may efficiently learn complex traffic patterns, advance real-time adaptability and effectively handle high-dimensional state spaces. But the approach might often struggle with uncertainty during training and might need large amount of training datasets and tuning for convergence. Subsequently, in MARL methods various agents are implemented at multiple connections and for optimize network-wide traffic flow every coordinating and learning is connected with each other. Consequently, this technique allows decentralized control and improves scalability for making it appropriate for wide-range of city

Authorized licensed use limited to: Nitte Meenakshi Institute of Technology. Downloaded on June 25,2026 at 06:58:27 UTC from IEEE Xplore. Restrictions apply.

traffic conditions. Yet, the model struggles with coordination between multiple agents which might be challenging.

- The Grouped Deep Q-Network (GDQN) converts the continuous traffic parameters into discrete congestion levels such as low, medium and high, which significantly improves learning stability, convergences speed and state space complexity when compared to existing DQN methods.
- A dynamic reward function is used to real time traffic conditions to adjust the wight include traffic density, waiting time and queue length, allows the more flexible, efficient and context-aware decisions under numerous traffic scenarios.
- The multi-agent traffic control system with limited agent communication, improves the global traffic flows, improves the cooperation among intersections and decision making under the traffic scenarios.
- Although grouping and reward shaping exist, this research incorporates state grouping with adaptive reward and multi-agent coordination, enhancing the convergence and scalability jointly. This combined development delivers the practical enhancement over standalone methods.

The organization of this research is structured as follows: the related works along with their advantages and disadvantages are given in section 2, overall working of proposed methodology is in section 3, experimental results with respective discussion is given in section 4 and the conclusion of this research is discussed in section 5 respectively.

II. LITERATURE

R. Shashidhar et al. [13] developed IoT and Computer Vision (CV) based Intelligent Road Lane Detection System (IRLDS). However, accidents have been increasing gradually over time at an alarming rate, which is main cause due to lack of driver attention. In particular image processing techniques include Hough transform, line extrapolation and edge detection to detect the road lanes very easily and accurate. Moreover, vision based and IoT able to reach real time performance in tracking and detection of road boundaries. Therefore, Hough transform used to find relative lines regarding right and left lane boundaries. Still, lack of performance measures and test methodologies, leads to trouble in road and loan detection.

Prashant Kumar Shukla et al. [14] suggested Modified Recursive Learning (MRL) for Traffic flow monitoring in software-defined network. Yet, anomaly detection models don’t do modification of traffic flow, which leads to poor recognition performance due to lack of traffic information. In addition, modified recursive learning focus on users that how to match the flow of traffic, which increases the transparency from traffic congestion. Subsequently, matching traffic flow control the policy to obtain the optimized traffic data for abnormalities detection. Simultaneously, model achieved accurate anomaly detection in traffic flow monitoring. Still, computational resource causes the degree of latency.

Jaime Villa et al. [15] developed Deep Learning and Edge Computing for Intelligent Infrastructure for Traffic Monitoring. In addition, Deep Sort improved the capabilities of tracking by integrating the information by deep learning. Moreover, speed estimation technique suitable for edge computing and efficient transformation technique provides the better geolocation while reducing the computational resource in edge computing. Therefore, accuracy of classification and localization was performed successfully most of the scenarios. However, model gets less accuracy in occlusion.

Mohammed Qader Kheder et al. [16] suggested IoT-aided robotics and deep learning techniques for Real-time traffic monitoring system (IoTRM). However, large number of vehicles on road has made the traffic regulations challenging, particularly in crowded and large cities. Subsequently, IoT road monitoring employed the cameras and sensors to collect road data to observe the vehicles in real time. Moreover, Autonomous Robotic Car (ARC) operated through integration of IoT cameras and sensors. Therefore, by applying the real analysis of data, the system improved in navigation and management. Although, ARC efficiently navigate through marked roads, but cannot navigate through unmarked roads.

Kerang Cao et al. [17] demonstrated Deep Reinforcement Learning (DRL) for Optimization Control of Adaptive Traffic Signal. However, existing traffic management lacks the adaptability and complex traffic conditions. In addition, DRL used the action selection strategy to train faster, also the number of vehicles is effectively reduced by the verification. Moreover, DRL uses the state and reward and enhance the Convolutional Neural Network (CNN) so that reflect the state of traffic intersection more efficiently. Moreover, DRL model outperformed the existing model in traffic signal optimization. However, model has the poor steadiness and slow.

III. METHODOLOGY

In order to carry out this research, a Multi-Agent (MA) system is applied to collect real-time traffic parameter data from a simulated environment like information on queue length, the number of vehicles, waiting time, and traffic density. A combined objective is to reduce the congestion, waiting time, and queue length using a weighted optimization function, confirming balanced traffic flow development. These parameters will be fed to a deep reinforcement learning agent to determine the MA action space per intersection like changing the traffic signal phase, extending the green light duration which leads to enhanced traffic efficiency. Additionally, a GDQN is employed to convert the continuous time series traffic parameters into discrete levels for the purpose of reducing the overall complexity of the state space and expediting respective learning process for the agent. Additionally, the reward optimization method has been implemented in order to permit the agent to adaptively determine reward values positive or negative based upon traffic conditions, permitting for more effective and context-aware decision-making to occur at the MA agent. Furthermore, the coordination will occur among neighboring MA agents through the partial sharing of information, permitting for greater levels of global encompassed optimization and scalability. Overall, the MA agent in the proposed framework achieved greater accuracy and efficiency than traditional methods. Future research efforts will include employing real-world datasets, utilizing advanced computer vision models (i.e., YOLO) for live traffic detection, and improving communication strategies among MA agents in order to support greater resiliency and real-time deployment in smart cities. The representation of the proposed GDQN-ARO model is shown in Fig 1.

Authorized licensed use limited to: Nitte Meenakshi Institute of Technology. Downloaded on June 25,2026 at 06:58:27 UTC from IEEE Xplore. Restrictions apply.

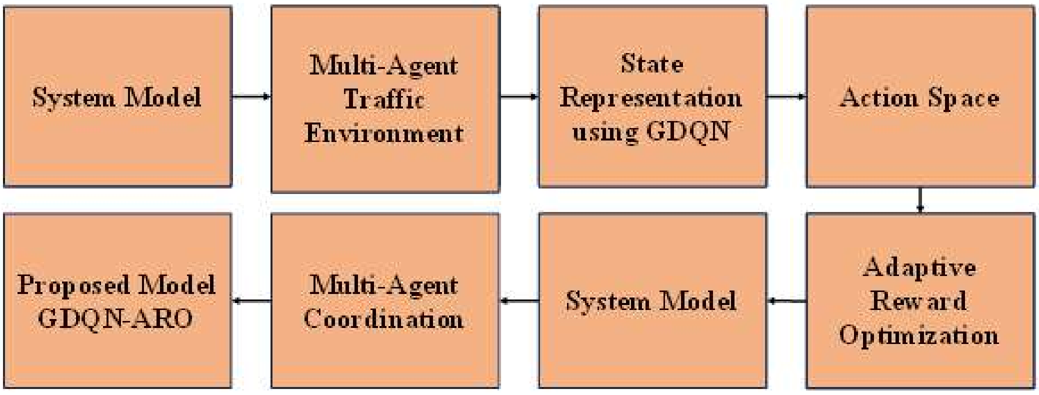

Fig. 1. The overflow diagram of proposed GDQN-ARO model

A. Multi-Agent Traffic Environment

In multi-agent traffic environment is demonstrated as the multi-agent system, where every traffic signal is represented as the individual and independent agent. Moreover, every agent makes the decisions and engage with its environment local through the observed traffic conditions. In addition, traffic simulation platform simulates the environment, where road networks use the numerous interactions. Unlike the centralized methods, the multi-agent confirms robustness and scalability by allocating the process of decision-making agents.

Every agent observes the parameters of local traffic include waiting time, queue length and vehicle count. Further, neighbouring interactions of limited information is shared to allowable coordination across the agents. This cooperation enables to system to attain global optimization instead emphasis solely on conditions of local traffic. Therefore, state of every agent at t time step is demonstrated in Equation (1):

$$s_t = \{q_t, w_t, d_t\} \quad (1)$$

Where d_t demonstrates traffic density, q_t represent queue length and $w_t$represents waiting time.

B. State Representation using GDQN

To enhance the learning efficiency, the GDQN used for the state representation, existing DQN models uses the raw numerical values, which results of slower convergence and high dimensional state. Therefore, GDQN decreases uses the traffic state grouping into discrete categories include high, medium and low congestion levels to decrease the complexity. The grouping function G(.) converts the input values into grouped states is shown in Equation (2):

$$s'_t = G(s_t) \quad (2)$$

Thresholds for low, medium, high congestion are represented as empirically based on traffic density and queue length to confirm the consistent state discretization. The queue length demonstrates in Equation (3):

$$G(q_t) = \begin{cases} Low, & q_t < \theta_1 \\ Medium, & \theta_1 \leq q_t < \theta_2 \\ High, & q_t \geq \theta_2 \end{cases} \quad (3)$$

Where θ_2 and θ_1 are predefined thresholds. This grouped demonstration decreases the input space dimensionality and enables the neural network to learn the specific patterns, which resulted the enhanced performance and faster convergence.

C. Action Space

In action space, every agent selects the action from the predefined action space that regulate the traffic signal phases and action set is demonstrated in Equation (4):

$$A = \{a_1, a_2, a_3\} \quad (4)$$

Where, a_1 balance the present signal phase, a_2 denotes the switch to next phase and a_3 represents the extended duration of green signal. Moreover, each action explicitly maps to traffic signal operations (hold, switch, extend green) confirming clear policy learning.

Therefore, the agent selects the optimal action through the estimated Q-values produced through GDQN model. The goal to choice action that enhance traffic flow and minimize the congestion at every interaction.

D. Adaptive Reward Optimization

The proposed model main component is the adaptive reward mechanism, which adjust dynamically through real time traffic conditions. Therefore, unlike the fixed reward functions, the adaptive reward function reflects the numerous performances demonstrates include traffic density, waiting time and queue length. The reward function is demonstrated in Equation (5):

$$r_t = -(\alpha q_t + \beta w_t + \gamma d_t) \quad (5)$$

Where, α , β , and γ are the weighting factors the demonstrate every parameter significance. The waiting time and queue length results in the higher reward of less negative values, which encourages the agent to optimize the traffic conditions. Furthermore, the weight parameters are dynamically tuned based on traffic density to arrange critical congestion, enhancing the context-aware learning. In addition, to enhance the adaptability, the weights are adapted dynamically through the traffic density, which is shown in Equation (6):

$$\alpha = f(d_t), \beta = f(d_t) \quad (6)$$

This confirms that the model gives the main priority to serious conditions include the peak hour congestion.

Authorized licensed use limited to: Nitte Meenakshi Institute of Technology. Downloaded on June 25,2026 at 06:58:27 UTC from IEEE Xplore. Restrictions apply.

E. Learning Process using GDQN

The GDQN algorithm uses every agent to learn the optimal policies by using environmental interaction. Therefore, the cumulative discounted reward maximizes overtime and Q-value function is updated through the standard Q-learning rule in Equation (7):

$$Q(s,a) = Q(s,a) + \eta \left[r_t + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \tag{7}$$

Where, η denotes the learning rate, γ refers discount factor, s' represent the next state and a' shows the next action and Neural network approximates Q-values utilizing the parameters updated through the loss minimization, confirming stable convergence and policy learning. The neural network is the Q-function is shown in Equation (8):

$$Q(s,a;\theta) \tag{8}$$

where θ denotes the network parameters and the loss function is used to train network, is shown in Equation (9):

$$L(\theta) = \mathbb{E}\left[\left(r_t + \gamma \max_{a'} Q(s',a';\theta^-) - Q(s,a;\theta) \right)^2 \right] \tag{9}$$

Where, θ^- represent the network target parameters. The model iteratively updates its parameters to improve decision-making and minimize this loss.

F. Multi-Agent Coordination

To improve the global traffic optimization, agents are exchanging the imperfect information with its surrounding interaction. Therefore, coordination enables every agent to consider the actions on adjacent intersections and local traffic conditions. The shared information is demonstrated in Equation (10):

$$s_t^{global} = s_t + \sum_{i \in N} s_t^i \tag{10}$$

Where, N denotes the neighboring agents. Agents exchange local traffic summaries at fixed intervals, allowing the cooperative decision-making and global optimization. This cooperative mechanism assists in decreasing the improves overall system performance and traffic spillover effects. The proposed framework iteratively operates, where every agent observes the environment, the GDQN uses the grouped states such as select action, takes the reward and policy update through coordination and interaction, the system coverages towards the optimal traffic signal strategy. The multi-agent learning integration, adaptive reward optimization and grouped state demonstration confirms the superior performance, faster convergence when compared to existing traffic signal methods.

IV. Experiment Results

In this research, GDQN-ARO model used for intelligent traffic signal optimization system. Simulation environment, dataset, and traffic scenarios are defined to confirm the reproducibility and fair evaluation. In addition, software environment used in this research Intel i7, 16 GB RAM and windows 11. In particular, python libraries include NumPy, pandas and TensorFlow. The following equations used in this experiment results are shown in Equations (11) – (13):

$$Precision = \frac{TP}{TP+FP} \tag{11}$$

$$Recall = \frac{TP}{TP+FN} \tag{12}$$

$$F1\,Score = 2 * \frac{precision \times recall}{precision + recall} * 100 \tag{13}$$

Where, TP, TN refers true positive and true negative, FP refers false positive and FN refers false negative respectively.

A. Performance Analysis

The proposed GDQN-ARO compared with existing models include IRLDS, IoTRM and MRL. The proposed GDQN-ARO model is shown in Table 1.

TABLE I. PERFORMANCE ANALYSIS OF PROPOSED GDQN-ARO MODEL

Performance Analysis	Precision (%)	Recall (%)	F1-Score (%)
IRLDS	95.3	95.6	95.4
IoTRM	96.5	95.9	96.2
MRL	96.7	96.3	97.1
Proposed GDQN-ARO	99.5	98.6	98.8

From the Table 1, the proposed GDQN-ARO model outperformed all existing models such as IRLDS, IoTRM and MRL includes precision of 99.5%, recall of 98.6% and f1-score of 98.8%. Moreover, IRLDS contains precision of 95.3%, recall of 95.6% and f1-score 95.4%. In addition, IoTRM contains precision of 96.5%, recall of 95.9% and f1-score of 96.2%. Additionally, MRL contains precision of 96.7%, recall of 96.3% and f1-score of 97.1%. The GDQN-based state grouping mechanism decreases the continuous traffic data complexity by classifying it into discrete congestion levels such as low, medium and high. This simplification allows the improved stability, better generalization and improved stability compared to existing DQN models.

B. Comparative Analysis

The proposed GDQN-ARO compared with existing models include YOLOv5. The proposed GDQN-ARO model is shown in Table 2.

TABLE II. COMPARATIVE ANALYSIS OF PROPOSED GDQN-ARO MODEL

Comparative Analysis	Precision (%)	Recall (%)	F1-Score (%)
YOLOv5 [16]	98.4	97.5	-
Proposed GDQN-ARO	99.5	98.6	98.8

From the Table 2, the proposed GDQN-ARO model outperformed all existing models such as GDQN-ARO includes precision of 99.5%, recall of 98.6% and f1-score of 98.8%. Moreover, YOLOv5 precision of 98.4%, recall of 97.5%. YOLOv5 is included as a vision-only baseline for vehicle detection, while ITFM-CV-IoT represents a complete traffic management framework. The GDQN-based state grouping mechanism decreases the continuous traffic data complexity by classifying it into discrete congestion levels such as low, medium and high. This simplification allows the improved stability, better generalization and improved stability compared to existing DQN models.

Authorized licensed use limited to: Nitte Meenakshi Institute of Technology. Downloaded on June 25,2026 at 06:58:27 UTC from IEEE Xplore. Restrictions apply.

C. Traffic Scenario-Based Analysis

To assess the proposed GDQN-ARO model effectiveness, the experiments are tested in the three conditions such as low, medium and heavy congestion tasks. Therefore, these scenarios are simulated in queue lengths, arrival rates and vehicle density.

In low traffic scenario, the model effectively manages the traffic with smooth vehicle and minimal waiting time by preventing unreliable signal changes. In medium traffic scenario, GDQN-ARO model dynamically adapts to moderate the congestion through optimizing signal timing based on the practical traffic parameters to enhance throughput and reducing queue length.

In heavy traffic scenario of peak hours, the system involves in severe congestion and high vehicle density. The proposed GDQN-ARO model uses the adaptive reward optimization for efficient critical intersections, which dynamically adjusts according to the traffic conditions. Moreover, GDQN grouping mechanism decreases the state complexity, allows the stable decision making and faster even in dense traffic environments.

In total, the results represent the GDQN-ARO model balance the consistent performance with significantly improve the traffic conditions, highlights the suitability and robustness for practical smart city applications.

D. Discussion

In this research, an intelligent traffic signal optimization system using GDQN-based multi-agent systems with adaptive reward optimization is proposed to improve the urban traffic flow and permit more efficient real-time decision-making at signal-controlled intersections. However, existing methods, especially standard DQNs and multi-agent reinforcement learning systems, are not ideal for signalized intersection traffic management because of their deficiencies like very high dimension state spaces, slow training convergence rates, lack of training stability, and poor inter-agent coordination. In addition, traditional methods have been controlled to utilizing the static reward functions that do not adapt to the continual changes occurring as a function of varying traffic volumes, with most of these methods offering poor performance through peak hours and under various traffic densities. These challenges highly limit the effectiveness of signalized intersections within urban areas; therefore, presenting challenges related to scalability, effectiveness of learning within a signals-based system, and real-time applicability of the learning attained from previous interactions in compound urban environments. To overcome these challenges, the proposed approach is implementing a GDQN-based grouping of states which applies various levels of representation of traffic conditions such as low congestion, medium congestion, and high congestion to highly reduce the complexity of the state space through the aggregation of states thus providing quicker training convergence and improved stability of training. Moreover, the proposed method will employ an adaptive reward optimization method, which adapts the weight of the reward in relation to real-time traffic conditions like queue length, wait time(s), and density. This permits the utilization of context-sensitive and responsive decision-making procedures to manage traffic at signal-controlled intersections. Through the utilization of multi-agent coordination with limited information sharing, the proposed method will progress the cooperation between intersections and thereby solve issues related to decentralized control and increase global traffic efficiency. Thus, providing a more efficient alternative to traditional and current DQ-based methods, facilitating faster convergence, enhanced scalability and higher levels of overall traffic efficiency when implemented within a smart city application.

V. Conclusion

The primary aim of this research is to establish a highly effective and intelligent traffic signal optimization system that applies DRL techniques to help decrease congestion, reduce waiting times, and increase the complete flow of traffic within urban areas. In order to carry out this research, a Multi-Agent (MA) system is applied to collect real-time traffic parameter data from a simulated environment like information on queue length, the number of vehicles, waiting time, and traffic density. These parameters will be fed to a deep reinforcement learning agent to determine the MA action space per intersection like changing the traffic signal phase, extending the green light duration which leads to enhanced traffic efficiency. Additionally, a GDQN is employed to convert the continuous time series traffic parameters into discrete levels for the purpose of reducing the overall complexity of the state space and expediting respective learning process for the agent. Additionally, the reward optimization method has been implemented in order to permit the agent to adaptively determine reward values positive or negative based upon traffic conditions, permitting for more effective and context-aware decision-making to occur at the MA agent. Furthermore, the coordination will occur among neighboring MA agents through the partial sharing of information, permitting for greater levels of global encompassed optimization and scalability. Overall, the MA agent in the proposed framework achieved greater accuracy and efficiency than traditional methods. Future research efforts will include employing real-world datasets, utilizing advanced computer vision models (i.e., YOLO) for live traffic detection, and improving communication strategies among MA agents in order to support greater resiliency and real-time deployment in smart cities.

References

[1] Shabab, K.R., Ali, S.M. and Zaki, M.H., 2023. Deep reinforcement learning-based short-term traffic signal optimizing using disaggregated vehicle data. Data science for transportation, 5(2), p.13.

[2] Bálint, K., Tamás, T. and Tamás, B., 2022. Deep reinforcement learning based approach for traffic signal control. Transportation Research Procedia, 62, pp.278-285.

[3] Ijjina, E.P. and Mohan, C.K., 2014, December. Facial expression recognition using kinect depth sensor and convolutional neural networks. In 2014 13th International Conference on Machine Learning and Applications (pp. 392-396). IEEE.

[4] Haddad, T.A., Hedjazi, D. and Aouag, S., 2022. A deep reinforcement learning-based cooperative approach for multi-intersection traffic signal control. Engineering Applications of Artificial Intelligence, 114, p.105019.

[5] Wang, T., Zhu, Z., Zhang, J., Tian, J. and Zhang, W., 2024. A large-scale traffic signal control algorithm based on multi-layer graph deep reinforcement learning. Transportation Research Part C: Emerging Technologies, 162, p.104582.

[6] Zhao, R., Luk, W., Niu, X., Shi, H. and Wang, H., 2017, July. Hardware acceleration for machine learning. In 2017 IEEE computer society annual symposium on VLSI (ISVLSI) (pp. 645-650). IEEE.

[7] Mao, F., Li, Z. and Li, L., 2022. A comparison of deep reinforcement learning models for isolated traffic signal control. IEEE Intelligent Transportation Systems Magazine, 15(1), pp.160-180.

Authorized licensed use limited to: Nitte Meenakshi Institute of Technology. Downloaded on June 25,2026 at 06:58:27 UTC from IEEE Xplore. Restrictions apply.

[8] Li, J., Wu, X., Xu, M. and Liu, Y., 2022. Deep reinforcement learning and reward shaping based eco-driving control for automated HEVs among signalized intersections. Energy, 251, p.123924.

[9] Wu, J., Song, Z. and Lv, C., 2023. Deep reinforcement learning-based energy-efficient decision-making for autonomous electric vehicle in dynamic traffic environments. IEEE Transactions on Transportation Electrification, 10(1), pp.875-887.

[10] Tunc, I. and Soylemez, M.T., 2023. Fuzzy logic and deep Q learning based control for traffic lights. Alexandria Engineering Journal, 67, pp.343-359.

[11] Wang, C., Xu, Y., Zhang, J. and Ran, B., 2022. Integrated traffic control for freeway recurrent bottleneck based on deep reinforcement learning. IEEE Transactions on Intelligent Transportation Systems, 23(9), pp.15522-15535.

[12] Pang, A., Wang, M., Pun, M.O., Chen, C.S. and Xiong, X., 2026. iLLM-TSC: Integration reinforcement learning and large language model for traffic signal control policy improvement. IEEE Transactions on Vehicular Technology.

[13] Shashidhar, R., Arunakumari, B.N., Manjunath, A.S., Ahuja, N.J., Hoang, V.T., Tran-Trung, K. and Belay, A., 2022. Computer Vision and the IoT-Based Intelligent Road Lane Detection System. Mathematical Problems in Engineering, 2022(1), p.4755113.

[14] Shukla, P.K., Maheshwary, P., Subramanian, E.K., Shilpa, V.J. and Varma, P.R.K., 2023. Traffic flow monitoring in software-defined network using modified recursive learning. Physical Communication, 57, p.101997.

[15] Villa, J., García, F., Jover, R., Martínez, V. and Armingol, J.M., 2024. Intelligent infrastructure for traffic monitoring based on deep learning and edge computing. Journal of Advanced Transportation, 2024(1), p.3679014.

[16] Kheder, M.Q. and Mohammed, A.A., 2024. Real-time traffic monitoring system using IoT-aided robotics and deep learning techniques. Kuwait Journal of Science, 51(1), p.100153.

[17] Cao, K., Wang, L., Zhang, S., Duan, L., Jiang, G., Sfarra, S., Zhang, H. and Jung, H., 2024. Optimization control of adaptive traffic signal with deep reinforcement learning. Electronics, 13(1), p.198.

Authorized licensed use limited to: Nitte Meenakshi Institute of Technology. Downloaded on June 25,2026 at 06:58:27 UTC from IEEE Xplore. Restrictions apply.