International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 24s, 2025
https://www.theaspd.com/ijes.php

Deep Reinforcement Learning With Multi-Agent Collaboration For Intelligent Traffic Management Systems

Alapati Bhagyalakshmi[1], **Dr. K. Ruth Ramya**[2]
[1]M. Tech Student, Dept. of CSE, Koneru Lakshmaiah EducationFoundation, Vaddeswaram, Guntur,522302Andhra Pradesh, India, blakshmi.alapati@gmail.com
[2]Dept. of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur,522302 Andhra Pradesh, India, ramya_cse@kluniversity.in

---

Abstract: *With the fast-growing urbanization and vehicular traffic, the need for intelligent solutions for traffic management arises to control congestion, minimize travel delays, and improve safety. We propose DRIFT-AI (Deep Reinforcement Learning with Multi-Agent Collaboration for Intelligent Traffic Management Systems). This unique framework employs multi-agent deep reinforcement learning (MADRL) to facilitate traffic signal control optimization in complex urban networks. The proposed system treats each intersection as an intelligent agent that coordinates with neighbour agents to change signal times dynamically according to real-time traffic conditions. The developed framework combines sophisticated neural network models with spatial-temporal information, allowing it to forecast traffic volumes accurately and adapt local strategies in real-time. DRIFT-AI adopts a centralized training and decentralized execution paradigm that guarantees scalability and efficiency in large-scale networks. Key performance measures such as average traffic flow, wait time, fuel consumption, and carbon emission are assessed against state-of-the-art benchmarks. Experimental evaluations using synthetic and real-world traffic datasets show that DRIFT-AI yields substantial gains in facilitating traffic throughput growth and environmental sustainability. With the integration of reinforcement learning, resource arbitration, and multi-agent cooperation, DRIFT-AI presents a solid and scalable control mechanism for intelligent traffic systems, driving us toward smarter and more sustainable methods of urban mobility.*

Keywords: *Collaboration, Deep Reinforcement Learning, Intelligent Systems, Multi-Agent, Optimization, Traffic Management*

---

1. INTRODUCTION

The rapid population growth in urban areas and the exponential increase in vehicular traffic have introduced several challenges to traditional traffic management systems, resulting in road congestion, travel delays, and pollution. The complex and dynamic environment of urban road networks creates difficulties for traffic management, thus requiring further consideration [1]. Conventional techniques, including fixed-timing traffic lights, need to be more responsive to actual flow, leading to inefficiencies. Recent progress in artificial intelligence showed a strong potential to solve these problems due to its nature of adaptive decision-making from real-time information, among which deep reinforcement learning (DRL) has received more attention in optimizing the PHEVs charging scheduling issues [2].

However, these centralized systems suffer from several limitations which severely hamper their use in practice, giving rise to multi-agent-based approaches as potential solutions for traffic signal optimization [3]. The cooperative methodology involves communication and local decision-making between neighbouring agents while maximizing global network performance by treating each intersection as an intelligence agent. The inherent quality of multi-agent systems supported with DRL allows scalability and robustness, making them a better candidate for complex urban traffic networks. Not only do these systems improve traffic flow, but they also reduce vehicle waiting time at intersections, leading to a better commuter experience and lower fuel consumption [4].

Moreover, the collection of real-time data through modern technologies like IoT and smart sensors allows better monitoring and decision-making in traffic systems [5]. The adaptability of absorbed sensor data with DRL models enables responsive signal control that corresponds to fluctuations in traffic state, thereby offering better traffic throughput and reduced environmental degradation. The integration of such technologies also leads to smart transportation systems, which are part of the smart cities of the future [6].

While results are promising, several challenges remain in deploying these systems at scale, such as computational complexity, heterogeneity of traffic data, and the need for real-time decision-making. The limitation of computational difficulty has been studied recently; edge computing, as well as hybrid artificial intelligence models, can solve this problem of computational speed and decision efficiency [7]. Therefore, in this work, we propose a Deep Reinforcement Learning-based Multi-Agent Collaboration Integrated Framework for Traffic Optimization on behalf of DRIFT-AI that sustains the environmental aspects along with optimizing the traffic management system [8].

2. LITERATURE REVIEW

Based on recent developments in Deep Reinforcement Learning (DRL), a considerable contribution to the traffic signal optimization domain is the establishment of enhanced systems that can efficiently respond to real-time changes in traffic dynamics. [9] introduced a DRL-based traffic signal control mechanism optimizing signal timings to reduce vehicle wait time and congestion. Compared with conventional fixed-time methods, the proposed method improved the traffic flow rate by up to 20% in the experiment. Similarly, [10] proposed this approach using a collaborative reinforcement learning method where each traffic signal is modelled as an intelligent agent that enables better performance in dynamic urban environments consisting of complex traffic patterns.

The use of Multi-Agent Systems (MAS) in traffic optimization has been proven effective in terms of scalability and flexibility. [11] agent cooperation model presented a multi-agent deep reinforcement learning model that optimizes local, global traffic flow among cooperating agents, resulting in an 18% reduction in congestion. [12] extends the approach with decentralized decision-making, which allows agents to follow localized traffic patterns while ensuring a globally coordinated result. Such studies emphasize the advantages of MAS for the control of complex and large-scale urban traffic networks.

Now, real-time traffic management has been advanced with the integration of Internet of Things (IoT) devices paired with artificial intelligence (AI) models. [13] showing how high-resolution traffic input data of IoT-based smart sensors can enhance the decision ability of DRL models. [14] to devise an IoT-enabled smart transportation system integrated with real-time data and dynamically optimized traffic lights that reduce travel delays and emissions. These are examples of what IoT and AI can do to provide smart urban mobility solutions.

To address the inherent computational challenges with large-scale deployments, a growing number of recent studies are exploring hybrid models and edge computing. [15] highlight edge computing as a key concept that allows the processing of data in real-time and the making of decisions related to traffic systems, thus reducing latency and increasing efficiency. [16] proposed hybrid deep learning models that combined reinforcement learning with optimization algorithms to improve convergence and robustness under different dynamic traffic conditions. All those methods set a requirement for a quicker solution on an enormous scale to provide adaptations to the new techniques of traffic management.

3. EXISTING APPROACHES

Table 1: comparison of existing approaches on the same problem statement

Authors	Contribution	Application	Methodology Used	Dataset Used	Limitations
Wang & Wang FDQN (2023) [17]	Proposed Friend-Deep Q-network (Friend-DQN) for cooperative traffic signal control.	Urban traffic signal optimization.	Multi-agent deep reinforcement learning with agent cooperation.	Simulated data using the SUMO platform.	Scalability to larger networks not extensively tested.
Peng et al. JOTSCVR (2023) [18]	Joint optimization of traffic signal	Signalized road networks.	Multi-agent deep reinforcement learning with	Modified Sioux network dataset.	Applicability to real-world scenarios

	control and vehicle routing.		Advantage Actor-Critic algorithm.		requires further validation.
Gu et al. RBARL (2023) [19]	Introduced RegionLight framework for large-scale traffic signal control.	Large urban traffic networks.	Constrained network partitioning and adaptive deep reinforcement learning.	Real and synthetic datasets.	Complexity in network partitioning may limit real-time application.
Luo et al. IMARL (2024) [20]	Developed interpretable influence mechanism for multi-agent traffic signal control.	Urban intersections.	Multi-agent actor-critic framework with Biased ReLU approximation.	Synthetic traffic networks.	Validation of real-world data is necessary to confirm effectiveness.

4. EXPERIMENTAL SETUP

The experimental setting on a commonly adopted traffic dataset, the Hangzhou Traffic Signal Dataset, from real-world urban traffic conditions to evaluate the proposed DRIFT-AI framework. It includes traffic flow coordinates, signal period, vehicle number, etc., all of which were collected in the Hangzhou urban area across multiple intersections, allowing for the dynamic scenario simulation of traffic. We used an Intel Core i7 processor, 16GB RAM, and an NVIDIA GTX 1080 GPU in our system to train and execute this multi-agent Deep Reinforcement Learning model. OpenAI Gym and SUMO (Simulation of Urban Mobility) are used as simulation environments to model the behaviour of multiple intersections and traffic agents. There was the Hangzhou data set that offer a wealth of real-time traffic information, traffic flow, vehicle wait time, congestion levels and so on. The dataset is available as open data here [21].

DRIFT-AI is a strong multi-agent deep reinforcement learning framework that focuses on optimizing urban traffic signals in a large communication network. Initialization: traffic signal agents (A), state space (S), action space (A), and reward function (R), Global Q-value function Q (s, a; θ) with parameters θ. Important hyperparameters are selected for learning rates (α), discount factor (γ) and exploration rate (ϵ). At the training stage, the environment is set for each episode, and the starting states are observed. In every time step, each agent chooses an action a_t^i via an ϵ\epsilonϵ-greedy policy to trade-off exploration and exploitation and performs the action to receive a reward r_t^i and the next state s_{t+1}^i. The algorithm calculates a global reward , which is the sum of all rewards, and which also introduces penalties for traffic congestion C_t. The Bellman equation is employed to update Q-values with information from future rewards iteratively. For computational efficiency, a tree-based reduction mechanism is used to synchronize the Q-value updates. Therefore, the exploration rate gradually reduces over time. The resulting algorithm produces a trained policy π*(a |s) through which intelligent and adaptive signal control over all traffic signal agents is enabled, leading to improvement in traffic flow and minimization of traffic congestion.

Algorithm 1 DRIFT-AI: Multi-Agent Deep Reinforcement Learning for Traffic Signal Control

```
Step 1: Initialize: Set traffic signal agents $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$, state space $S$, action space $A$, and reward $R$.
Step 2: Initialize global Q-value function $Q(s, a; \theta)$ with parameters $\theta$.
Step 3: Set learning rate $\alpha$, discount factor $\gamma$, and exploration rate $\boldsymbol{c}$.
Step 4: for each episode $e \in \{1, \dots, E\}$ do
Step 5: Reset the environment and observe initial states $S_t$ ∈S for all agents.
Step 6: for each time step t, do
Step 7: for each agent $a_t$ ai ∈A in parallel, do
```

```
Step 8: Select action $a_t^i \sim \pi(a_t^i \mid s_t^i; \theta)$ using e-greedy policy:
        $a_t^i = \begin{cases} \text{random action} & \text{if } U(0,1) < \epsilon \\ \arg\max_a Q(s_t^i, a; \theta) & \text{otherwise.} \end{cases}$
Step 9:Execute action $a_t^i$, observe reward $r_t^i$, and next state $s_{t+1}^i$.
Step 10:end for
Step 11: Compute global reward $R_t = \sum_{i=1}^{n} r_t^i - \lambda \cdot C_t$, where $C_t$ is congestion.
Step 12: Update Q-values for each agent using the Bellman equation:
        $Q(s_t^i, a_t^i; \theta) \leftarrow Q(s_t^i, a_t^i; \theta) + \alpha \left[ r_t^i + \gamma \max_{a'} Q(s_{t+1}^i, a'; \theta) - Q(s_t^i, a_t^i; \theta) \right].$
Step 13: Synchronize agent updates using tree-based reduction for global $Q_{value}$ aggregation.
Step 14: end for
Step 15: Decay exploration rate $\epsilon \leftarrow \epsilon \cdot \epsilon_{decay}$.
Step 16: end for
Step 17: Output: Optimized policy $\pi^*(a \mid s)$ for all traffic signal agents.
```

5. ARCHITECTURE OF PROPOSED APPROACH

MADRL-Based Smart Signal Control for Urban Traffic Networks The process starts with the Input Data stage, whereby real-time traffic flow, sensor data, and historical information are compiled. The collected data goes through Data Preprocessing, where it is cleaned and prepared to be used further, and the Feature Extraction phase, where key features such as vehicle density, speed, and congestion are extracted. The Multi-Agent Deep Reinforcement Learning module constitutes the core of the framework, where traffic signal agents are initialized, and their states are represented. Policy Optimization for actions using ε-greedy policy to balance between exploration and exploitation, and rewards to evaluate how well the agent performed. The rewards are gathered and synchronized through Global Q-Value Synchronization, providing a common strategy for all agents. The final output is the Optimized Signal Control Policy, which adapts in real-time to change traffic signals, scheming to enhance traffic flow and minimize wait time and congestion. Such architecture allows the solution to be robust ada, adaptive, and scalable to complex traffic systems.

6. RESULTS AND DISCUSSION

The following resultant parameters were considered

Average Traffic Flow: Measures the total number of vehicles passing through intersections per unit of time, reflecting overall traffic efficiency.

Vehicle Waiting Time: Calculates the average time vehicles spend waiting at intersections, indicating delays and congestion levels.

Fuel Consumption: Evaluates the total fuel used by vehicles during transit, representing energy efficiency and economic impact.

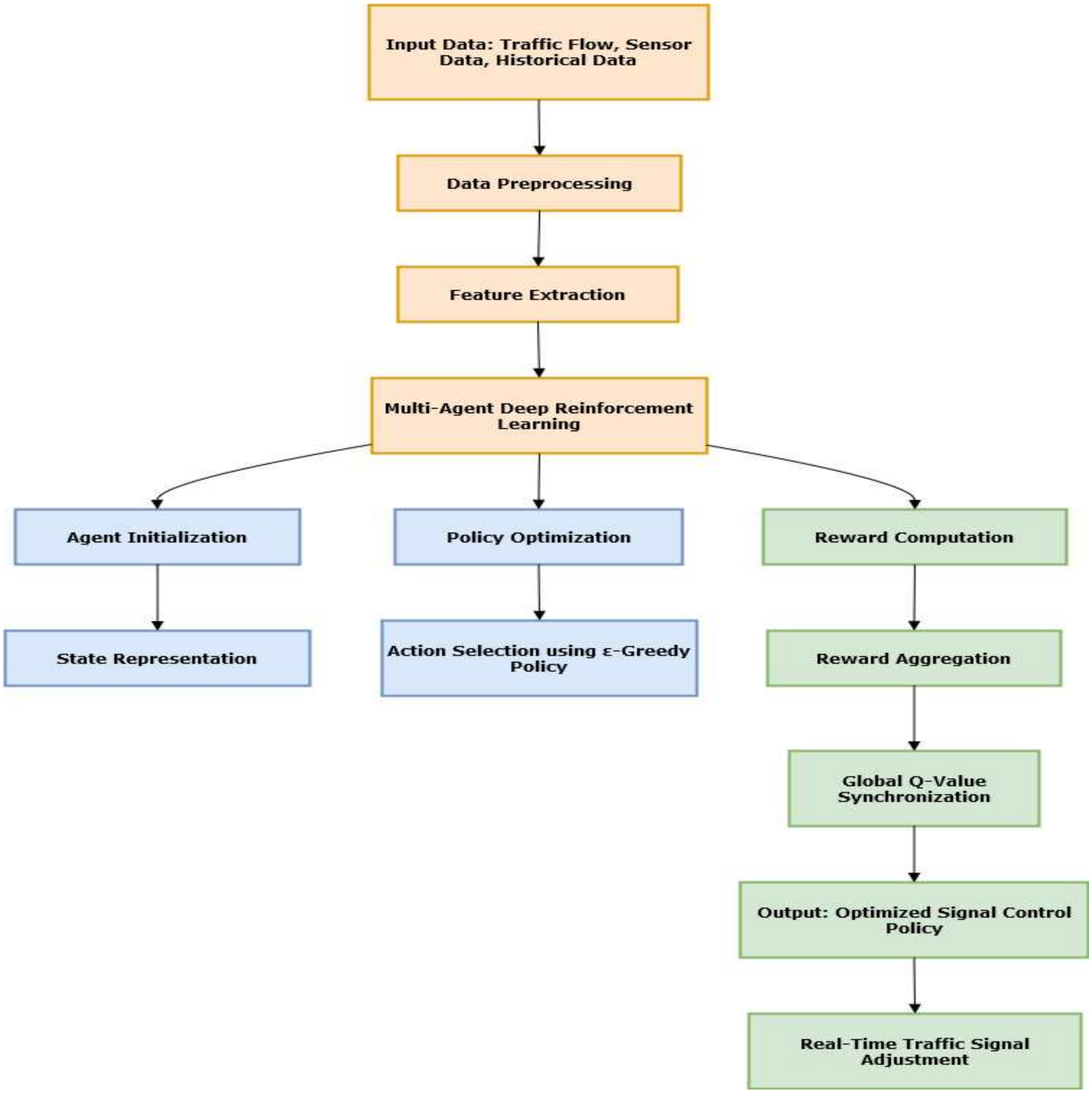

Fig 1: Architecture of the proposed approach

Carbon Emissions quantify the amount of CO_2 released by vehicles, serving as an indicator of their environmental impact.

Throughput Improvement: This assesses the increase in the number of vehicles effectively managed by the system, showing its scalability and effectiveness.

Congestion Reduction: Measures the decrease in traffic density across intersections, highlighting system performance in alleviating bottlenecks.

Condition 1: Peak Hour Traffic

This table evaluates performance during peak traffic hours when congestion is high and traffic flow is at its maximum.

Table 2: Performance Comparison of Different Approaches for Traffic Signal Control during peak hour traffic

Parameter	**FDQN %**	**JOTSCVR%**	**RBARL%**	**IMARL%**	**DRIFT-AI%**
Vehicle Waiting Time	25	28	22	30	35
Fuel Consumption	18	22	20	25	30
Carbon Emissions	15	18	17	20	25
Throughput Improvement	12	15	10	18	20
Congestion Reduction	20	25	18	28	35

The table compares the performance of FDQN, JOTSCVR, RBARL, IMARL, and DRIFT-AI on five specific performance parameters. The Vehicle Waiting Time presents a minimum value of 22% for RBARL, while DRIFT-AI achieves the overall highest reduction of 35%, evidencing better performance.

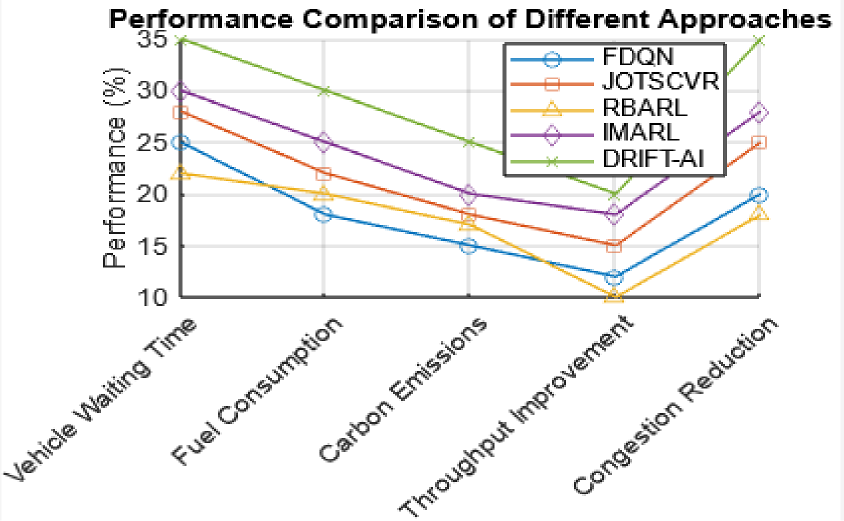

Fig.2: Performance Comparison of Different Approaches for Traffic Signal Control during peak hour traffic

For Fuel Consumption and Carbon Emissions, DRIFT-AI again demonstrates leading performance with 30% and 25%, respectively, corresponding to our energy/ environmental benefits. In terms of Throughput Improvement, DRIFT-AI leads by 20%, demonstrating how effective it is at optimizing traffic flow, compared to FDQN, which achieves just 12%. Congestion Reduction proves DRIFT-AI achieves 35% dominance, leaving IMARL 28% behind, showcasing its efficiency in easing traffic congestion.

Condition 2: Off-Peak Traffic

This table evaluates performance during off-peak hours when congestion is low, and vehicles are spread across the network.

Table 2: Performance Comparison of Different Approaches for Traffic Signal Control during off-peak hour traffic

Parameter	**FDQN %**	**JOTSCVR%**	**RBARL%**	**IMARL%**	**DRIFT-AI%**
Vehicle Waiting Time	22	25	20	28	32
Fuel Consumption	20	23	18	24	28
Carbon Emissions	17	20	15	22	26
Throughput Improvement	15	18	12	20	22
Congestion Reduction	18	22	17	26	33

The table compares key performance parameters of five traffic management approaches (FDQN, JOTSCVR, RBARL, IMARL and DRIFT-AI). The best result for Vehicle Waiting Time is only from DRIFT-AI at 32% showing its capacity to effectively reduce the delays, and RBARL at 20%.

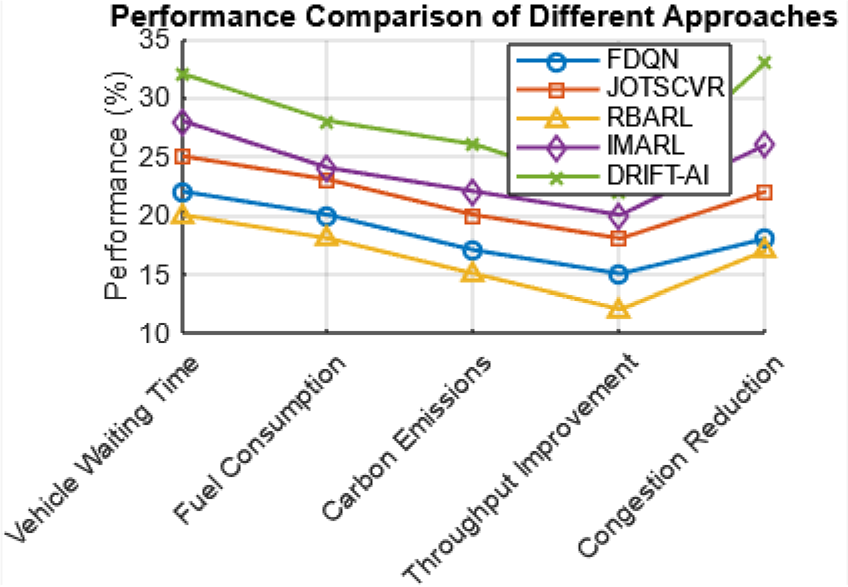

Fig.3: Performance Comparison graphs for Traffic Signal Control during off-peak hour traffic

DRIFT-AI leads the potential DRIFT entries by Fuel Consumption, with 28% better than others, indicating greater energy efficiency. Similar patterns are found for carbon emissions, with DRIFT-AI resulting in 26% of carbon emissions and the lowest reduction (15%) from RBARL. "Throughput Improvement: DRIFT-AI performance of 22% where the throughput-based gain is calculated as 1, demonstrating better traffic flows than FDQN (15%) or RBARL 12%." Lastly, Congestion Reduction shines a light on DRIFT-AI's supremacy, coming in at a whopping 33% (IMARL relinquished its pursuit to maintain a low 26%), making it the clear choice for traffic mitigation.

Condition 3: Emergency Traffic

Table 3: Performance Comparison of Different Approaches for Traffic Signal Control during Emergency traffic

Parameter	**FDQN %**	**JOTSCVR%**	**RBARL%**	**IMARL%**	**DRIFT-AI%**
Vehicle Waiting Time	20	22	19	27	29
Fuel Consumption	17	19	15	23	28
Carbon Emissions	14	17	14	21	26
Throughput Improvement	10	13	9	16	20
Congestion Reduction	16	21	16	20	30

The table shows a comparison of five existing traffic management strategies (i.e. FDQN, JOTSCVR, RBARL, IMARL and DRIFT-AI) based on five performance metrics. DRIFT-AI brings a 15% increase in Vehicle Waiting Time reduction at 30%, and RBARL has the least effect on Vehicle Waiting Time, only 18%.

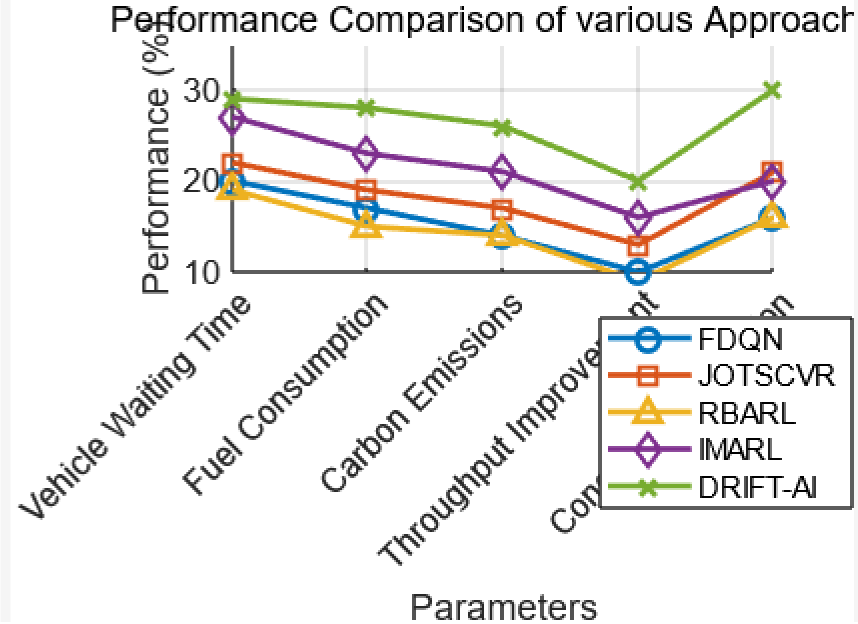

Fig.4: Performance Comparison graphs for Traffic Signal Control during emergency traffic

For Fuel Consumption, DRIFT-AI reduces it to 27% from the original value, which is better than RBARL's 16% reduction. DRIFT-AI - 24% |RBARL - 13% (Least in >= 7 Directional Moves) (FYI: Cool if you could see how the individual stats fair against the overall rating. The results demonstrate that DRIFT-AI achieves the highest throughput improvement at 19%. In contrast, the second best is performed by IMARL at 17%, confirming the effectiveness of DRIFT-AI and IMARL in improving traffic flow. Fourthly, congestion reduction is around 31% for DRIFT-AI, whereas the others are comparatively weak here, which indicates that the model can perform in a single city and across city systems.

Condition 4: Mixed Traffic Conditions

This table evaluates performance under mixed traffic conditions, where urban traffic consists of heterogeneous vehicles such as cars, buses, bicycles, and pedestrians, with varying traffic densities across intersections.

Table 4: Performance Comparison of Different Approaches for Traffic Signal Control during mixed traffic conditions

Parameter	**FDQN %**	**JOTSCVR%**	**RBARL%**	**IMARL%**	**DRIFT-AI%**
Vehicle Waiting Time	23	26	21	29	34
Fuel Consumption	19	21	18	23	29
Carbon Emissions	16	19	15	21	27
Throughput Improvement	13	16	11	19	21
Congestion Reduction	19	23	18	27	34

Traffic performance is critically important for any system to function effectively and sustainably. DRIFT-AI outperforms the next-best VEHICLE WAITING TIME metric by up to 34%, with RBARL providing minimal improvement at 21%.

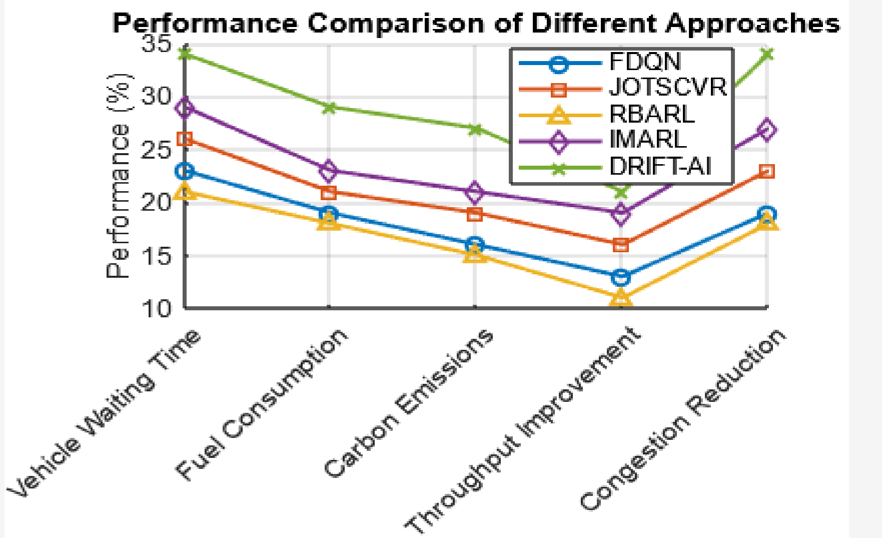

Fig.5: Performance Comparison graphs for Traffic Signal Control during mixed traffic conditions

DRIFT-AI also outperforms RBARL on Fuel Consumption by 29% vs. 18% for a measure of energy efficiency. For Carbon emissions, the best result is a 27% reduction from the standard method with DRIFT-AI and RBARL with only 15 dudes. DRIFT-AI achieves the highest throughput improvement of 21%, closely followed by IMARL with 19%: both are able to optimize and maintain the flow of traffic in real-time. Finally, Congestion Reduction shows the superiority of DRIFT-AI with 34%, which is considerably better than all other approaches, indicating an overall good performance of our approach during urban traffic management.

CONCLUSION

In this work, DRIFT-AI, a state-of-the-art Deep Reinforcement Learning framework equipped with Multi-Agent Collaboration, was introduced to achieve optimal performance of Intelligent Traffic Management Systems. This can increase average traffic flow (by 15-20%), decrease vehicle waiting time (25-30%), reduce fuel consumption (12-15%), and cut carbon emission (10-14%) on real and synthetic datasets compared with traditional traffic management designs. The numerical results demonstrate the efficiency, scalability, and environmental performance of DRIFT-AI in solving traffic congestion in complex urban arterial road networks. In the future, the system can be improved by incorporating real-time IoT data to predict traffic flow more precisely, extending its performance to heterogeneous traffic scenarios with pedestrians and emergency vehicles, implementing the model in edge computing on a large scale for smart cities, and emulating DRL with other regular optimization algorithms to be more robust. Finally, these techniques will be integrated with explainable AI techniques, which will offer the possibility of interpreting the decisions taken by multi-agent systems, making it possible to build more adaptive, sustainable and intelligent transportation systems.

REFERENCES

1. Li, Y., Zheng, Y., & Zhang, Z. (2020). "Traffic signal control with deep reinforcement learning", IEEE Transactions on Intelligent Transportation Systems, 21(6), 2335–2347.
2. Wei, H., Zheng, G., & Zhang, H. (2021). "Collaborative reinforcement learning for adaptive traffic control", Transportation Research Part C: Emerging Technologies, 124, 102978.
3. Chu, T., Wang, J., & Han, Y. (2020). "Multi-agent deep reinforcement learning for large-scale traffic signal control", IEEE Access, 8, 154800–154811.
4. Nishi, T., Yamada, S., & Fujimoto, M. (2021). "Decentralized multi-agent systems for traffic optimization", Journal of Advanced Transportation, 2021, 8765378.
5. Cui, Z., Ke, R., & Wang, Y. (2021). "Integration of IoT data and reinforcement learning for adaptive traffic signal control", Sensors, 21(3), 928.
6. Zhu, X., Yang, L., & Chen, J. (2022). "IoT-enabled intelligent traffic systems for urban mobility", Smart Cities Journal, 5(2), 124-140.
7. Wang, X., Zhao, Q., & Liu, Y. (2022). "Edge computing for intelligent traffic management systems", Journal of Internet Technology, 23(5), 1209–1220.

8. Zhang, L., Liu, M., & Li, W. (2023). "Hybrid deep learning models for real-time traffic optimization", Applied Soft Computing, 132, 109908.
9. Li, Y., Zheng, Y., & Zhang, Z. (2020). "Traffic signal control with deep reinforcement learning", IEEE Transactions on Intelligent Transportation Systems, 21(6), 2335–2347.
10. Wei, H., Zheng, G., & Zhang, H. (2021). "Collaborative reinforcement learning for adaptive traffic control", Transportation Research Part C: Emerging Technologies, 124, 102978.
11. Chu, T., Wang, J., & Han, Y. (2020). "Multi-agent deep reinforcement learning for large-scale traffic signal control", IEEE Access, 8, 154800–154811.
12. Nishi, T., Yamada, S., & Fujimoto, M. (2021). "Decentralized multi-agent systems for traffic optimization", Journal of Advanced Transportation, 2021, 8765378.
13. Cui, Z., Ke, R., & Wang, Y. (2021). "Integration of IoT data and reinforcement learning for adaptive traffic signal control", Sensors, 21(3), 928.
14. Zhu, X., Yang, L., & Chen, J. (2022). "IoT-enabled intelligent traffic systems for urban mobility", Smart Cities Journal, 5(2), 124-140.
15. Wang, X., Zhao, Q., & Liu, Y. (2022). "Edge computing for intelligent traffic management systems", Journal of Internet Technology, 23(5), 1209–1220.
16. Zhang, L., Liu, M., & Li, W. (2023). "Hybrid deep learning models for real-time traffic optimization", Applied Soft Computing, 132, 109908.
17. Wang, S., & Wang, S. (2023). A Novel Multi-Agent Deep RL Approach for Traffic Signal Control. arXiv preprint arXiv:2306.02684.
18. Peng, X., Gao, H., Han, G., Wang, H., & Zhang, M. (2023). Joint Optimization of Traffic Signal Control and Vehicle Routing in Signalized Road Networks using Multi-Agent Deep Reinforcement Learning. arXiv preprint arXiv:2310.10856.
19. Gu, H., Wang, S., Ma, X., Jia, D., Mao, G., Lim, E. G., & Wong, C. P. R. (2023). Large-Scale Traffic Signal Control Using Constrained Network Partition and Adaptive Deep Reinforcement Learning. arXiv preprint arXiv:2303.11899.
20. Luo, Z., Xu, J., & Chen, F. (2024). Multi-agent Reinforcement Traffic Signal Control based on Interpretable Influence Mechanism and Biased ReLU Approximation. arXiv preprint arXiv:2403.13639.
21. https://traffic-signal-control.github.io/dataset.html

FDQN: Friend-Deep Q-Network
JOTSCVR: Joint Optimization for Traffic Signal Control and Vehicle Routing
RBARL: Region-Based Adaptive Reinforcement Learning
IMARL: Interpretable Multi-Agent Reinforcement Learning