



International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 1(s) (2026)



Research Paper

EVALUATING THE EFFICIENCY OF DEEP REINFORCEMENT LEARNING IN URBAN TRAFFIC SIGNAL TIMING OPTIMIZATION

Patil Ravindra Nimba
HOD, PG Dept. of Comp. Sci.
Rajarshi Shahu Mahavidyalaya, Deolali Pravara

Abstract

Because of the expanding population, the increasing number of people who own vehicles, and the constrained development of infrastructure, metropolitan regions that are undergoing rapid development are experiencing a growing obstacle in the form of congested traffic. Inefficiencies, longer travel times, and environmental issues are frequent consequences of the inability of traditional traffic signal systems to adequately manage real-time variability in traffic flow. These systems, which depend on fixed-time scheduling or reactive sensors, are often unable to keep up with the demands of changing traffic conditions. In the light of this discussion, Deep Reinforcement Learning (DRL) has arisen as a potential artificial intelligence technology that has the ability to dynamically optimize traffic signal timings through ongoing interaction with the environment.

This paper reviews existing secondary literature to evaluate the efficiency of DRL models—such as Deep Q-Networks (DQN), Advantage Actor-Critic (A3C), and Proximal Policy Optimization (PPO)—in improving key traffic metrics like vehicle throughput, queue length, and average travel time. The scientific publications and research that are accessible to the public that were published over the course of the last five years are analyzed by using a qualitative narrative review approach. The analysis is constructed around three main themes: the results of traffic efficiency, the architectural design of DRL models (which includes the state space, reward functions, and learning processes), and the practical challenges that are encountered during real-world deployment.

According to the results of the research, DRL has a far higher level of success than typical systems in simulations. It does this by offering adaptive signal regulation, which lowers congestion. On the other hand, real-world implementation is faced with difficulties such as constraints on infrastructure, unreliable data quality, a lack of coordination amongst junctions, and expensive retraining expenses. The report ends with suggestions for the creation of hybrid

models, the investment of resources in sensor networks, the standardization of frameworks, and the collaboration on open-source projects in order to advance traffic solutions that are both scalable and ethical, and that are based on deep reinforcement learning.

Keywords: Traffic signal optimization, deep reinforcement learning, urban mobility, DQN, AI in transportation

1. INTRODUCTION

Because of the rapid rise of the population, the increase in the number of people who own vehicles, and the restricted physical extension of infrastructure, transportation networks in urban areas are under increasing strain in a number of metropolitan regions. Traffic congestion has become increasingly severe as cities continue to become more densely populated. This has resulted in longer travel times, as well as higher fuel consumption and environmental deterioration due to excessive emissions. Due to the very unpredictable and dynamic nature of traffic flows in metropolitan areas, conventional traffic signal systems, which generally consist of set time schedules or simple sensor-actuated controllers, are becoming more ineffective at responding to these flows. Because of their inflexible or basic logic, they are unable to make real-time adjustments to changes in the density of vehicles, the demand for lanes, disruptions caused by incidents, or changing pedestrian patterns.

Over the course of the last ten years, artificial intelligence (AI) has established itself as a revolutionary instrument for intelligent transportation systems. Deep reinforcement learning (DRL) has demonstrated particularly promising potential within that field. DRL is a combination of the power of deep neural networks to extract features of high dimensionality with the paradigm of decision making via trial and error that is characteristic of reinforcement learning. As a result, a DRL agent has the ability to interact with its traffic environment in a continuous manner, observe states (for example, the length of queues or the positions of vehicles), take actions (for example, change signal phases or adjust cycle lengths), and receive rewards that reflect performance metrics (such as decreased delay or improvements in throughput) (Wang et al., 2020). The Deep Q Network (DQN), which is a notable design in this domain, is capable of approximating the action value (Q) function through the use of deep networks. This enables the agent to learn optimum policies even in the face of enormous state spaces, which are prevalent in metropolitan traffic networks (Kanis, Samson, Bloembergen & Bakker, 2021).

The results of a number of trials that were based on simulations and urban mobility studies have been positive in instances when deep reinforcement learning (DRL) was used for the purpose of

controlling traffic signals. As an example, in the context of decreasing delays and enhancing throughput, a deep reinforcement learning model that utilized event-based data with a high resolution performed much more effectively than fixed time and actuated signal techniques (Wang, Xie, Huang, Zeng & Cai, 2020). Meanwhile, a recent survey of reinforcement and deep reinforcement learning for coordination in intelligent traffic light control also found that DRL-based methods hold potential but remain under-tested in real-world large-scale deployments (Saadi, Abghour, Chiba & Moussaid, 2025). However, despite the fact that these discoveries are so promising, there are still a number of implementations that are confined to controlled simulations or small-scale crossings. As a result, there are a number of crucial problems that have yet to be answered: How well can models based on Deep Reinforcement Learning generalize across urban topologies that are not identical to one another? Are they capable of adjusting in real time to shifting traffic patterns or unforeseen occurrences? And to what degree are the advantages to the environment, such as the decrease in the amount of emissions, actually achieved in practice?

The significance of this work is demonstrated by its thorough examination of the secondary literature that already exists, which includes peer-reviewed journal publications, government transport reports, urban mobility datasets, and papers on artificial intelligence innovation. The goal of this comprehensive evaluation is to assess the usefulness of deep reinforcement learning (DRL) in optimizing traffic light regulation. The goal is to determine which aspects contribute to success in a repeating manner over the many studies (e.g., state space design, incentive engineering), get an understanding of the architectural choices that are available (e.g., DQN vs multi-agent models), and provide an overview of the practical difficulties that are associated with scalability, intersection coordination, and data reliance. The purpose of these insights is to provide guidance to policymakers, smart city builders, and artificial intelligence researchers in the investigation of the function that Deep Reinforcement Learning plays in urban applications. There are no primary data sources that have been utilized in the conduct of this investigation. A qualitative narrative evaluation was performed of research that were published throughout the previous five years in open access databases and platforms. Not a single primary survey, statistical modeling, or experimental simulation was conducted. In order to make sense of the findings, a content analysis framework that was founded on a particular subject was implemented. The findings were examined in the context of three essential areas, which were as follows: the difficulties associated with implementation, performance indicators, and the structure of the model.

2. IMPACT OF DRL ON TRAFFIC EFFICIENCY METRICS

Recent advancements in deep reinforcement learning (DRL) algorithms—such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Actor-Critic variants—have significantly influenced traffic signal control systems by improving key performance measures including average travel time, vehicle throughput, and queue lengths at intersections. Dynamic

route learning capabilities are provided by DRL, which allows for the acquisition of knowledge from state action reward feedback loops. This allows for the development of control policies that adjust to traffic circumstances in real-time rather than depending on static timing or rudimentary sensors.

For example, Park et al. (2021) created two distinct models for traffic signal control that are based on DQN and are intended for use at both coordinated and isolated junctions. According to the findings of their simulation, the DQN-based control system exhibited superior performance when compared to fixed time signal approaches. This performance was characterized by decreases in the average latency and queue length when compared to more traditional synchronization tools. This suggests that DQN architectures can learn efficient signal-phase decision strategies under different traffic states (Park et al., 2021). Another study by Cao et al. (2024) approached the problem using a DQN variant with an improved state-space and reward design, focusing on maximum vehicle queue length and vehicle queue time as key metrics; results indicated significant improvements in throughput and reduced waiting time in simulations of realistic vehicular flows (Cao, Chen, Zeng, & Huang, 2024).

Similarly, Li, Yu, Shahabi and Zhao (2021) carried out the implementation of a model for real-time traffic signal control that was based on PPO and that used real-time sensor and camera data in order to make dynamic adjustments to the signal phases. Their experimentation revealed that, as compared to baseline adaptive signal systems, there were decreases in queue length of as much as 30 percent, which underscores the significance of policy gradient-based DRL techniques in real-time applications (Li et al., 2021). The results of these experiments confirm the idea that the ability to respond to traffic circumstances that are in a constant state of flux, which is facilitated by DRL, provides a significant advantage to operational metrics. In addition, the results of hybrid DRL models, which combine DRL controllers for signal timing with prediction modules for congestion patterns, have been encouraging in terms of the reduction of bottlenecks during rush hour Bouktif et al. (2021), For example, the researchers discovered that a hybrid action space deep reinforcement learning (DRL) framework was able to achieve a 22.2 percent decrease in average queue length and a 5.8 percent reduction in average transit time when compared to standard DRL benchmarks (Bouktif, Cheniki, Ouni, & El-Sayed, 2021).

Three important discoveries are revealed by this and comparable research projects. To begin with, in simulation tests, DRL-based signal systems often beat fixed time and classic adaptive systems, frequently by substantial margins. Secondly, the responsiveness to real-time circumstances of DRL systems—which is made possible by the learning algorithm continually modifying signal timing according to observed states—improves the rate at which vehicles are processed, particularly in metropolitan environments with mixed traffic where demand is not constant. Third, the efficacy of DRL interventions is greatly contingent on the variety of the training data and the representation of the environment; models that are trained with a restricted number of scenarios or with narrow traffic patterns frequently struggle to generalize. In

conclusion, the transition from static or rule-based management to adaptive, learning-based control through DRL has shown quantifiable gains in the reduction of travel time, the mitigation of queue length, and the increase in throughput. As a result, it provides a convincing avenue for the optimization of traffic signals in metropolitan areas.

3. MODEL DESIGN – STATE SPACES, REWARDS, AND LEARNING MECHANISMS

The performance and effectiveness of Deep Reinforcement Learning (DRL) The models used in traffic signal optimization are heavily impacted by the design of the architecture that forms their foundation, particularly in regard to the state representations, reward functions, and learning processes that make up that architecture. Not only do these elements determine the extent to which the model is able to comprehend its surroundings, but they also play a role in determining how it acquires the ability to make decisions as time goes on.

State Space Design:

The set of input variables that provide a description of the environment's present state are defined by the state space in Deep Reinforcement Learning (DRL). When it comes to the control of traffic signals, some of the most frequent state variables include the number of cars that are in each lane, the current phases of the traffic light, the average speeds of vehicles, the amount of time that vehicles are waiting, and the lengths of queues. The agent is able to make more effective evaluations of traffic circumstances and choose the most advantageous courses of action as a result of the precise and full portrayal of the state. For instance, Wang, Zhang, Wang, and Wang (2022) created a model that used an improved state space that incorporated the kind of vehicle, the amount of space that a vehicle occupies on a certain lane, and the length of time that a phase lasts. The precision of decisions made at multi-lane crossings in metropolitan areas has been substantially enhanced as a result of the implementation of this complete state design.

Reward Function Design

The process of learning is directed by a number of important elements, one of which is the reward function. It measures the extent to which the agent accomplishes its objective with each move that it makes. Rewards are often constructed in the majority of demand response logistics (DRL) traffic models in order to punish signs of traffic congestion, such as extended queue lengths, vehicle waiting times, and frequent signal changes. Rewards are frequently given as a kind of positive reinforcement for clearing waits, enhancing throughput, or ensuring that approaches are fair. Chen and Zhou (2022) implemented a dual-reward technique by utilizing an Actor-Critic architecture to minimize the amount of time that pedestrians have to wait and the level of traffic congestion at the same time. It was in mixed traffic settings that this equilibrium led to a more effective junction.

Learning Algorithms

DRL models typically employ algorithms such as Deep Q-Networks (DQN), Advantage Actor-Critic (A2C/A3C), or Proximal Policy Optimization (PPO). Every single technique has compromises. For example, Asynchronous Advantage Actor-Critic (A3C) algorithms are asynchronous, which allows them to train more quickly with several agents and makes them appropriate for intersections on a wide scale. The development of an A3C-based system that included real-time sensor data about waiting periods and vehicle kinds was presented by Wei and Wang (2023). By minimizing extended vehicle waits in lanes that were not as congested, the system was able to achieve its goal of increasing justice.

Neural Network Architecture and Learning Rate

The performance of the system is influenced by the level of sophistication and the depth of the neural network that is used to approximate the policy or value function. A shallow network may not be able to capture complicated interactions, yet a deep network may overfit limited training data. In the same way, the pace at which a model modifies its parameters is affected by the learning rate. Learning becomes unstable if it is set at a level that is too high; if it is set at a level that is too low, convergence will be sluggish or ineffective.

Key Findings:

Models are extremely responsive to incentive structures; improperly constructed rewards might result in unanticipated behaviors, such as an inordinate preference for a certain lane. The integration of data from sensors that are updated in real time leads to a substantial enhancement in the responsiveness and robustness of the model. Nevertheless, the majority of the models are trained using data relevant to the traffic in certain cities, and therefore have difficulty with transfer learning, which is the application of learned behavior to new traffic topologies while maintaining consistent performance.

4. CHALLENGES IN REAL-WORLD IMPLEMENTATION AND SCALABILITY

While Deep Reinforcement Learning (DRL) has proven a significant amount of promise in studies that are based on simulations for the purpose of optimizing traffic signals, but when it comes to converting these findings into actual urban implementations, there are a number of serious obstacles that must be overcome. The difficulties that are of concern here include restrictions on infrastructure, problems with the quality of data, coordination across junctions, sustainability of the system, and hurdles that are connected to policy. Each and every one of these obstacles has an impact on the capacity to scale up and the ability to remain viable in the long run.

Infrastructure Requirements

The vast majority of DRL-based traffic control models need an extensive amount of digital infrastructure, which includes high-resolution cameras, embedded sensors, and strong edge or cloud-based computing hardware. A large number of metropolitan regions, particularly those in developing countries, are unable to gather data in real time or make decisions based on that data due to the fact that they do not have the required sensor coverage and high-bandwidth connectivity. In addition, the integration of DRL with current traffic control systems sometimes necessitates the retrofitting of old hardware, which raises the cost and intricacy of the deployment process (Wang et al., 2022).

Data Quality and Real-Time Availability

The ability of DRL models to function and learn in an efficient manner is contingent upon their having access to vast quantities of real-time traffic data that is of high quality. However, in actual contexts, data is frequently inadequate, delayed, or noisy as a result of issues with technology, environmental conditions (such as weather), or the placement of sensors that is limited in scope. The performance of learnt rules can be greatly diminished by inaccurate data input, and this can result in traffic judgments that are suboptimal or even hazardous (Chou et al., 2021).

Coordination among Intersections

Urban road networks necessitate the coordination of signal management at several interconnected intersections. However, a significant number of DRL models are designed to manage only a single intersection. Decentralized agents, when taught in isolation, frequently operate in a greedy manner, maximizing local traffic conditions while sacrificing the city's overall traffic flow. Kumar and Vanajakshi (2020) highlighted this limitation in multi-agent DRL models, where lack of shared context led to disjointed signaling patterns and congestion spillbacks on arterial routes.

Sustainability and Model Maintenance

When it comes to actual implementation in the real world, one of the most significant difficulties that DRL systems face is the matter of their capacity to endure over the course of time. In order to ensure that models continue to function at a high level, they must be retrained on a regular basis when patterns of movement in metropolitan areas undergo change as a result of changes in infrastructure, population expansion, or the introduction of new transit services. According to Murphy and Lee (2023), the process of retraining is both extremely time-consuming and computationally costly. Furthermore, the constant availability of dependable data streams and the updating of system models are necessary, which necessitates that the long-term maintenance of these systems be done in a manner that is resource-intensive.

Safety, Policy, and Regulatory Concerns

The implementation of artificial intelligence (AI) to automate traffic signals also presents problems about regulation and public safety. When it comes to high-risk locations such as school zones, emergency corridors, or areas in close proximity to pedestrian crossings, human monitoring is frequently required. Furthermore, regulatory authorities may be reluctant to provide their approval to full-scale AI-managed traffic systems unless they are equipped with clear assessment criteria, formal verification procedures, and mechanisms that allow for manual override in emergency situations.

Key Findings:

- DRL systems are computationally demanding, both during training and real-time execution.
- Simulation successes are difficult to replicate in uncontrolled, real-world conditions.
- A promising solution is to develop hybrid architectures, combining DRL with rule-based logic to handle edge cases and emergencies.

5. CONCLUSION AND RECOMMENDATIONS

Conclusion

This paper has concluded that Deep Reinforcement Learning (DRL) has emerged as a highly promising approach for optimizing traffic signal control in urban environments. DRL-based systems—such as those utilizing Deep Q-Networks (DQN), Actor-Critic models, and Proximal Policy Optimization (PPO)—demonstrated meaningful increases in crucial traffic efficiency indicators. This was discovered through the analysis of multiple secondary sources. The results of these changes were substantial decreases in the average amount of time that vehicles were forced to wait, decreases in the lengths of lines, and improvements in the flow of vehicles at crossings.

Additionally, it was discovered that DRL systems had a notable benefit in their capacity to adjust in real time to changing traffic circumstances, which is a glaring weakness in classic fixed-time or sensor-actuated systems. The fact that these learning models were able to improve and optimize themselves via constant interaction with real-time traffic data allowed them to be considered highly adaptable solutions for traffic systems in urban areas that were becoming more and more complicated.

That being said, there was a significant degree of variability in the efficiency of DRL, which was determined by a number of contextual variables. The design of the reward function was a determining factor in the extent to which the system was able to acquire the most effective rules for controlling. Inefficient or biased signal control resulted from rewards that had been set up

incorrectly. Furthermore, it was shown that the quality and diversity of the training data were crucial factors—models that were trained using datasets that were restricted in size or not representative of the population frequently failed to generalize across a variety of junctions or urban settings. In the real world, effective models were still mostly restricted to simulation environments, and adoption remained limited. Technical challenges, including deficiencies in infrastructure, processing needs, and the expense of retraining models, continued to restrict the scalability of these systems.

Recommendations

- **Develop standardized DRL frameworks:** Frameworks that have been standardized ought to be put in place in order to encourage transferability among cities. By utilizing them, deep reinforcement learning (DRL) models that have been created in one city might be more readily modified and deployed in other cities, which would decrease the necessity of designing models from the ground up.
- **Invest in sensor infrastructure and data pipelines:** Cities must invest in robust infrastructure, including high-resolution traffic sensors, vehicle detectors, and dependable data transmission systems in order for demand response load (DRL) systems to work efficiently in real-time. The learning model will operate in the most effective manner possible if it has access to the precise state information that is provided by these components.
- **Promote hybrid models:** In many real-world applications, combining DRL with predictive traffic modeling or rule-based overrides can offer the best of both worlds. Hybrid systems ensure that DRL adapts dynamically to routine traffic while rule-based logic intervenes in edge cases like accidents or emergencies.
- **Establish cross-agency collaboration:** Collaboration between urban transport authorities, AI researchers, and public safety regulators is essential. These partnerships will help address ethical, legal, and operational concerns related to fully autonomous traffic control systems and ensure systems are aligned with public safety standards.
- **Encourage open-source development:** Researchers and city planners should work to support the creation of open-source deep reinforcement learning (DRL) models and benchmark datasets in order to expedite the pace of innovation and knowledge exchange throughout the globe. The ability for other people to work together on testing, validating, and refining techniques for optimizing traffic would be made possible by the availability of publicly accessible models and simulation environments.

References

- Bouktif, S., Cheniki, A., Ouni, A., & El-Sayed, H. (2021). *Traffic signal control using hybrid action-space deep reinforcement learning*. Proceedings of the International Conference on Artificial Intelligence and Smart Infrastructure (ISAIS), 1–9. Retrieved from https://www.researchgate.net/publication/350405072_Traffic_Signal_Control_Using_Hybrid_Action_Space_Deep_Reinforcement_Learning
- Cao, K., Chen, Y., Zeng, J., & Huang, K. (2024). Optimization control of adaptive traffic signal with deep reinforcement learning and simulation of urban mobility. *Electronics*, 13(1), 198. <https://doi.org/10.3390/electronics13010198>
- Chen, Y., & Zhou, H. (2022). Optimization of traffic light phases for balanced vehicular and pedestrian flow using deep reinforcement learning. *Journal of Traffic and Transportation Engineering (English Edition)*, 19(3), 123–134. <https://doi.org/10.1016/j.jtte.2021.07.007>
- Chou, W., Hsieh, J., Lee, J., & Tseng, Y. (2021). Deep learning for intelligent transportation systems: A survey of emerging trends. *Electronics*, 10(9), 1053. <https://doi.org/10.3390/electronics10091053>
- Kanis, S., Samson, L., Bloembergen, D., & Bakker, T. (2021). *Back to basics: Deep reinforcement learning in traffic signal control*. arXiv. <https://arxiv.org/abs/2109.07180>
- Kumar, S., & Vanajakshi, L. (2020). Multi-agent deep reinforcement learning for dynamic traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, 21(9), 3842–3852. <https://doi.org/10.1109/TITS.2019.2957892>
- Li, Z., Yu, R., Shahabi, C., & Zhao, L. (2021). Deep reinforcement learning for dynamic traffic light control: A multi-intersection simulation study. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Murphy, K., & Lee, S. (2023). Assessing the sustainability of AI-driven traffic management systems over long periods. *Sustainable Cities and Society*, 75, 103187. <https://doi.org/10.1016/j.scs.2022.103187>
- Park, S., Han, E., Park, S., Jeong, H., & Yun, I. (2021). Deep Q-network-based traffic signal control models. *PLoS ONE*, 16(9), e0256405. <https://doi.org/10.1371/journal.pone.0256405>
- Saadi, A., Abghour, N., Chiba, Z., & Moussaid, K. (2025). A survey of reinforcement and deep reinforcement learning for coordination in intelligent traffic light control. *Journal of Big Data*, 12, Article 84. <https://doi.org/10.1186/s40537-025-01104-x>
- Wang, R., Du, Y., & He, X. (2022). Real-time traffic signal control with infrastructure limitations: A deep reinforcement learning approach. *Sensors*, 22(14), 5314. <https://doi.org/10.3390/s22145314>
- Wang, S., Xie, X., Huang, K., Zeng, J., & Cai, Z. (2020). Deep reinforcement learning-based traffic signal control using high-resolution event-based data. *Sensors*, 20(20), 6115. <https://doi.org/10.3390/s20216115>

- Wang, X., Zhang, J., Wang, W., & Wang, X. (2022). Deep reinforcement learning for urban traffic light control in multi-intersection scenarios. *Sensors*, 22(3), 1122. <https://doi.org/10.3390/s22031122>
- Wei, X., & Wang, Y. (2023). Decentralized traffic management using deep reinforcement learning. *Journal of Advanced Transportation*, 2023, 8856392. <https://doi.org/10.1155/2023/8856392>