Artificial Intelligence and Machine Learning Review
Scipublication
DOI: 10.69987/AIMLR.2024.50405

Adaptive Traffic Signal Timing Optimization Using Deep Reinforcement Learning in Urban Networks

Zhonghao Wu[1], Shaobo Wang[1.2], Chunhe Ni[2] , Jiang Wu[3]
[1] *Computer Engineering, New York University, NY, USA*
[1.2] *Computer Science and Engineering, Santa Clara University, CA, USA*
[2] *Computer Science, University of Texas at Dallas, Richardson, TX, USA*
[3] *Computer Science, University of Southern California, Los Angeles, CA, USA*
**Corresponding author E-mail: eliza6723s@gmail.com*

Keywords

Traffic Signal Control, Deep Reinforcement Learning, Deep Deterministic Policy Gradient, Adaptive Control Systems

Abstract

This document please leave the flashlight that uses the extra work (DRL) for local communication. LEFTMENT Activity GRARIAR (DDPG) MATAL-Multi-multi-multi-multisgre traffic IET, QUEUE LONGS, and Distance Models. Problems the management of the controller is designed to continue with the operating sites including reductions, by maintenance operations. The Network Network architecture is designed, featuring special components and mental health regulations from cars. Procedures are used and measured two test platform and real car information from the main city of metrozopolitan from 12 intersections. Experiment that the recommended suggestions have enhanced achievement processes existing, including the suspension of 23.5% in the average attack. The latest-world of validity is completed in a 6-month-monthly basis. Using the use of electric wiring for local work and clouds managed cooperation, improve the intelligence application.

Introduction

1.1 Research Background and Significance

Quick strategy and increase cars in cars have been causing serious trouble for the larger city for the larger cities today. City County Major City is relevant, fuel consumption, smoke, and public, and public quality. Traditional-time intersective-time and actuated issues are not able to do the best tracks, especially at the time without thinking about[1]. These restrictions are driven to the traffic lights that can be treated the traffic lights as a result of real time.

The emergence of deep reinforcement learning (DRL) has provided new opportunities to address complex traffic signal control problems. DRL combines the perception capabilities of deep neural networks with the decision-making abilities of reinforcement learning, enabling end-to-end learning of control policies from raw traffic data[2]. Unlike traditional optimization methods that rely on simplified traffic models and assumptions, DRL can learn optimal control strategies directly from interactions with the traffic environment, making it particularly suitable for adaptive traffic signal control in complex urban networks[3].

The significance of this research lies in several aspects: (1) The proposed adaptive traffic signal control system can significantly reduce average vehicle delay and improve traffic efficiency by learning optimal timing strategies from historical traffic patterns and real-time data[4]; (2) The DRL-based approach can handle complex traffic scenarios and uncertain conditions without requiring explicit traffic models; (3) The system can coordinate multiple intersections to achieve network-level optimization, leading to better overall traffic performance compared to isolated intersection control[5].

1.2 Research Status and Problems

Current research on traffic signal control can be broadly categorized into three approaches: fixed-time control, actuated control, and adaptive control. Fixed-time control uses predetermined signal timing plans based on historical traffic data but lacks flexibility to handle varying traffic conditions[6]. Actuated control adjusts signal timing based on real-time vehicle detection but typically follows simple rules that may not be optimal. Adaptive control systems can dynamically optimize signal timing parameters according to traffic conditions, representing the most promising direction for modern traffic management[7].

Recent advances in artificial intelligence have sparked interest in applying DRL to traffic signal control. Various DRL algorithms have been proposed, including Deep Q-Network (DQN), Deep Deterministic Policy Gradient (DDPG), and Proximal Policy Optimization (PPO)[8]. These approaches have demonstrated promising results in simulation environments. Nevertheless, several critical challenges remain unresolved: (1) The high-dimensional state and action spaces in traffic signal control make it difficult for DRL agents to learn effective policies[9]; (2) The complex interactions between multiple intersections pose challenges for coordinated control; (3) The gap between simulation and real-world implementation needs to be addressed[Error! Reference source not found.].

Existing studies have focused primarily on single intersection optimization or simplified network scenarios. The scalability and coordination issues in large-scale urban networks have not been adequately addressed. Additionally, most current approaches rely on complete traffic state information, which may not be available in practical applications due to limited sensor coverage and communication constraints[10]. The robustness and generalization capabilities of DRL-based traffic signal control systems under various traffic conditions and network configurations also require further investigation[11].

1.3 Research Content

This research proposes an adaptive traffic signal control system based on deep reinforcement learning for urban networks. The main research contents include:

The development of a comprehensive traffic signal optimization framework that integrates deep reinforcement learning with domain knowledge of traffic engineering[12]. The framework incorporates both microscopic traffic flow characteristics and macroscopic network-level coordination requirements. A novel state representation scheme is designed to capture essential traffic patterns while maintaining computational efficiency[13].

The design of an advanced DRL algorithm specifically tailored for traffic signal control. The algorithm employs a hierarchical architecture to decompose the complex control problem into manageable sub-tasks. A prioritized experience replay mechanism is implemented to improve learning efficiency by focusing on important traffic scenarios[14]. The reward function is carefully designed to balance multiple objectives including delay minimization, throughput maximization, and coordination maintenance.

A coordination mechanism is developed to enable collaborative control among multiple intersections. The mechanism utilizes graph neural networks to capture spatial dependencies between intersections and facilitate information sharing[15]. A decentralized learning approach is adopted to improve scalability while maintaining coordination through message passing between neighboring intersections. The system incorporates adaptive exploration strategies to balance exploitation of learned policies with exploration of new control strategies[16].

The research includes comprehensive experimental evaluation using both simulation platforms and real-world traffic data. The experiments investigate system performance under various traffic conditions, network configurations, and disturbance scenarios. Analysis of computational efficiency, convergence properties, and robustness under uncertainties is conducted. Comparative studies with existing methods demonstrate the advantages of the proposed approach in terms of delay reduction, throughput improvement, and coordination effectiveness[17].

2. Urban Traffic Signal Optimization Problem Modeling

2.1 Urban Traffic Network Environment Modeling

The urban traffic network is modeled as a directed graph $G = (V, E)$, where V represents the set of intersections and E denotes the set of road segments connecting these intersections. Each intersection $v \in V$ is characterized by its incoming and outgoing lanes, turning movements, and signal phases[18]. The road segments $e \in E$ are described by their length, capacity, and free-flow speed[19]. This graph representation captures both the topological structure and the traffic-related attributes of the urban network.

At each intersection, the traffic flow is controlled by a signal controller that operates multiple signal phases. A phase configuration P consists of a set of non-conflicting movements that can receive right-of-way simultaneously[20]. The phase sequence and duration determine the temporal allocation of green time to different traffic movements. Each movement is associated with a queue length q(t) and an arrival rate λ(t), which vary over time based on traffic demand patterns[21].

The traffic state at time t is represented by a comprehensive state vector s(t) that includes queue lengths, vehicle delays, traffic flow rates, and historical traffic patterns. Advanced detection systems provide real-time measurements of these traffic parameters through various sensors including loop detectors, cameras, and connected vehicle data. These measurements form the basis for adaptive signal control decisions[22].

2.2 Traffic Signal Timing Optimization Constraints

The optimization of traffic signal timing must satisfy multiple operational constraints to ensure safety and practicality. The minimum green time constraint ensures sufficient time for vehicles to safely clear the intersection: gi ≥ gmin, where gi is the green time for phase i and gmin is the minimum required green time. The maximum green time constraint prevents excessive delays to competing movements: gi ≤ gmax[23].

Phase sequence constraints maintain compatibility between different movements and prevent concurrent conflicting flows. The clearance time between phases must be sufficient to ensure safe transitions: yi + ri ≥ ymin + rmin, where yi and ri are the yellow and red clearance intervals for phase i**Error! Reference source not found.**. The cycle length C must fall within acceptable bounds: Cmin ≤ C ≤ Cmax to maintain coordination with adjacent intersections.

Coordination constraints ensure smooth traffic progression along arterial corridors. The offset θi between adjacent intersections must be properly set to facilitate green wave formation: θi = f(di,j, vi,j), where di,j is the distance between intersections i and j, and vi,j is the desired progression speed[24].

2.3 Mathematical Modeling and Objective Function

The traffic signal optimization problem is formulated as a multi-objective optimization problem. The primary objective is to minimize the total vehicle delay D across the network:

$$D = \sum t \sum i \sum m \ wi,m \times di,m(t)$$

where wi,m is the weight factor for movement m at intersection i, and di,m(t) is the corresponding vehicle delay at time t. The delay function incorporates both uniform and random delays based on traffic flow theory:

$$di,m(t) = d1,i,m(t) + d2,i,m(t)$$

where d1,i,m represents uniform delay assuming uniform arrivals, and d2,i,m accounts for random delay due to queue overflow.

Additional objectives include maximizing throughput T and minimizing the number of stops S:

$$T = \sum t \sum i \sum m \ fi,m(t)$$

$$S = \sum t \sum i \sum m \ si,m(t)$$

where fi,m(t) is the flow rate and si,m(t) is the number of stops for movement m at intersection i during time t.

2.4 Deep Reinforcement Learning Based Solution Framework

The traffic signal control problem is formulated as a Markov Decision Process (MDP) defined by the tuple (S, A, P, R), where S represents the state space, A is the action space, P is the state transition probability matrix, and R is the reward function[25]. The state space S encompasses traffic state variables including queue lengths, delays, and flow rates. The action space A consists of possible signal timing adjustments including phase splits, cycle lengths, and offsets.

The reward function R is designed to reflect multiple control objectives:

$$R(t) = w1 \times (-D(t)) + w2 \times T(t) + w3 \times (-S(t))$$

where w1, w2, and w3 are weight coefficients balancing different objectives. The negative signs before D(t) and S(t) convert the minimization objectives into maximization form.

The DRL agent learns a policy $\pi(a|s)$ that maps traffic states to optimal signal timing actions. The policy is represented by deep neural networks trained through interactions with the traffic environment. The value function Q(s,a) estimates the expected cumulative rewards of taking action a in state s:

$$Q(s,a) = E[\sum t\ \gamma t \times R(t)|s0 = s,\ a0 = a]$$

where γ is the discount factor balancing immediate and future rewards. The optimal policy maximizes the expected cumulative rewards across the entire network over the planning horizon.

3. Adaptive Traffic Signal Control Algorithm

3.1 Principles of Deep Reinforcement Learning

Deep Reinforcement Learning (DRL) combines deep neural networks with reinforcement learning principles to achieve end-to-end learning of control policies. In the context of traffic signal control, the DRL agent interacts with the traffic environment by observing traffic states and executing signal timing actions[26]. The learning process aims to maximize cumulative rewards through trial-and-error exploration and exploitation.

Table 1. Components of DRL-based Traffic Signal Control

Component	Description	Mathematical Representation
State Space	Traffic flow data, queue length	$s_t \in S$
Action Space	Signal timing parameters	$a_t \in A$
Reward Function	Delay, throughput metrics	$r_t = R(s_t, a_t)$
Value Function	Expected return estimation	$V(s_t) = E[\sum \gamma^k r_{t+k}]$
Policy Function	Control strategy	$\pi(a_t|s_t)$

The DRL framework employs both value-based and policy-based methods. Value-based methods learn the Q-function Q(s,a) to evaluate action-value pairs, while policy-based methods directly optimize the policy $\pi(a|s)$. The integration of deep neural networks enables the handling of high-dimensional continuous state and action spaces characteristic of traffic control problems.

3.2 State Space and Action Space Design

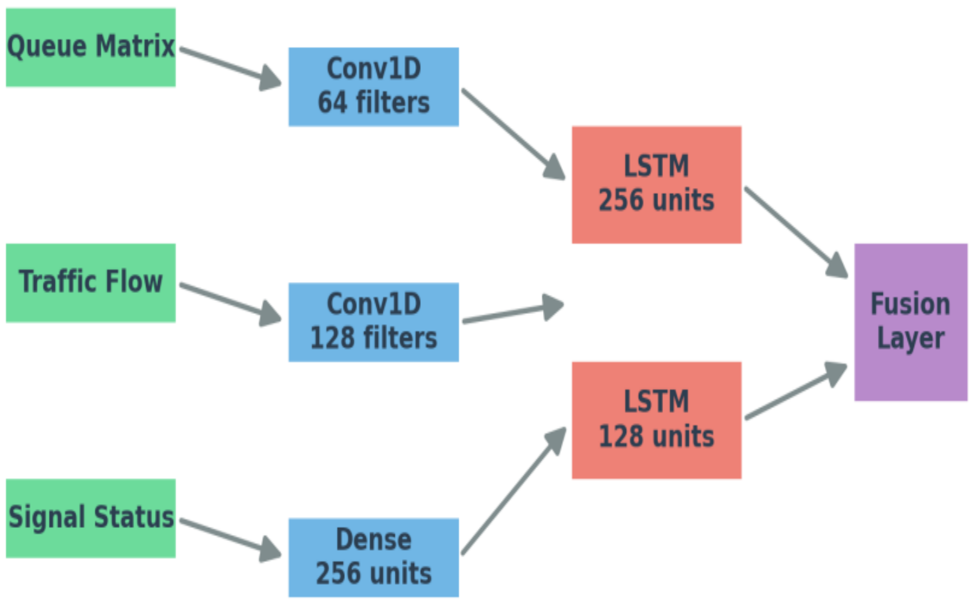
The state space design incorporates multiple traffic parameters to provide a comprehensive representation of intersection conditions. A multi-modal state representation scheme combines different data sources and temporal information.

Table 2. State Space Components and Dimensions

Component	Variables	Dimension	Update Frequency
Queue Length	Per lane queue	12 × 1	5s
Traffic Flow	Vehicle count	8 × 1	5s

Signal Status	Phase indicators	4 × 1	Real-time
Historical Data	Previous states	5 × 24	1h

Figure 1. Multi-modal State Representation Architecture



The figure demonstrates the hierarchical structure of state representation, featuring parallel processing branches for different data modalities. The architecture includes convolutional layers for spatial feature extraction from queue matrices, LSTM layers for temporal pattern learning from historical data, and fusion layers for integrated state representation.

The action space consists of adjustable signal timing parameters designed to balance control flexibility and computational tractability. The continuous action space enables fine-grained control adjustments while maintaining phase sequence constraints.

Table 3. Action Space Parameters

Parameter	**Range**	**Resolution**	**Update Rule**
Green Time	[15,60]s	1s	Continuous
Phase Split	[0.1,0.6]	0.01	Proportion
Cycle Length	[60,120]s	5s	Discrete
Offset	[-20,20]s	1s	Continuous

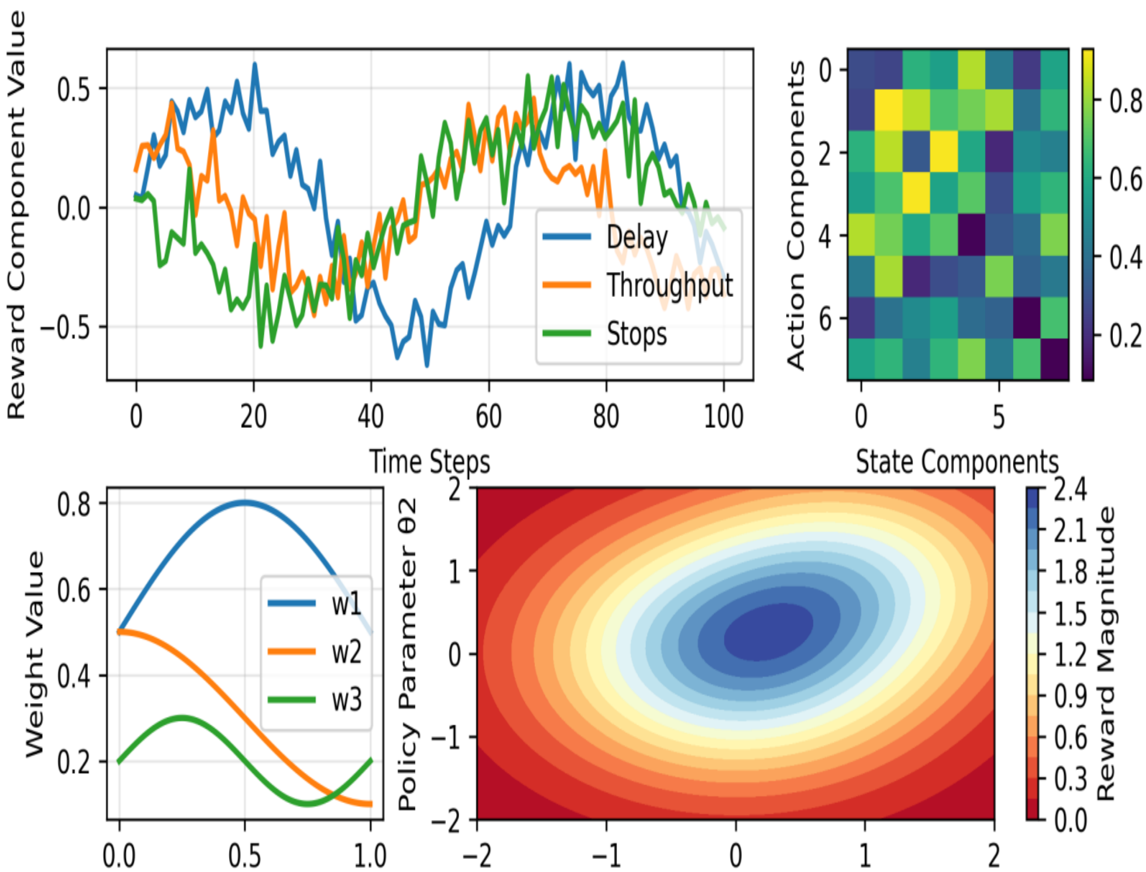
3.3 Reward Function Design and Optimization

The reward function incorporates multiple performance metrics weighted according to their relative importance in overall traffic optimization objectives.

Table 4. Reward Function Components and Weights

Metric	Weight	Calculation Method	Normalization
Delay	0.4	Queue-based	Min-max
Throughput	0.3	Flow-based	Softmax
Stops	0.2	Binary count	Exponential
Coordination	0.1	Phase difference	Linear

Figure 2. Reward Shaping and Optimization Process



The figure illustrates the reward calculation pipeline, including parallel computation of individual metrics, dynamic weight adjustment based on traffic conditions, and reward scaling mechanisms. The visualization incorporates heat maps for weight distributions and temporal evolution of reward components.

3.4 Deep Deterministic Policy Gradient Learning Algorithm

The DDPG algorithm implements an actor-critic architecture with deterministic policy gradient updates. The actor network $\mu(s|\theta\mu)$ generates deterministic actions, while the critic network $Q(s,a|\theta Q)$ evaluates action values[Error! Reference source not found.].

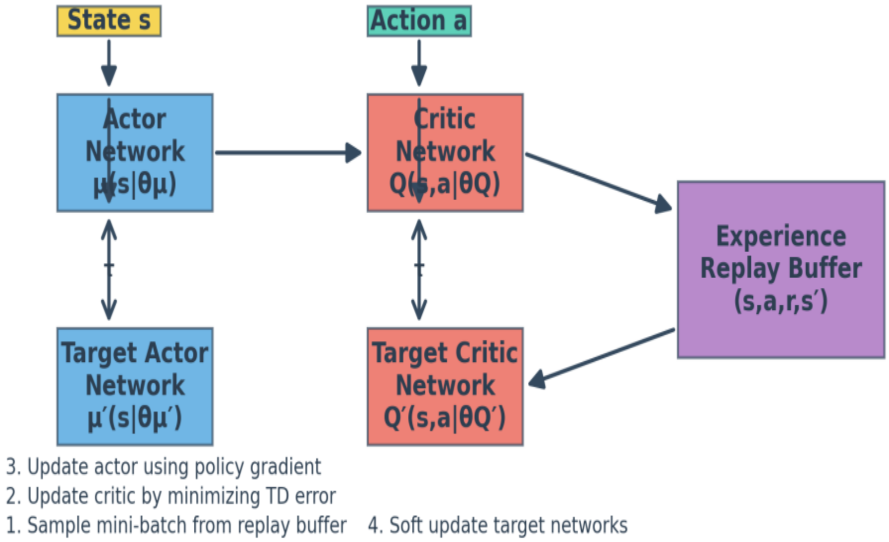
The critic is updated by minimizing the loss:

$$L(\theta Q) = E[(r + \gamma Q'(s',\mu'(s'|\theta\mu')|\theta Q')\ Q(s,a|\theta Q))^2]$$

The actor is updated using the deterministic policy gradient:

$$\nabla\theta\mu J \approx E[\nabla aQ(s,a|\theta Q)|a=\mu(s)\ \nabla\theta\mu\mu(s|\theta\mu)]$$

Figure 3. DDPG Architecture and Learning Process



The figure presents the dual network structure of DDPG, featuring parallel actor and critic networks with target network copies. The visualization includes gradient flow paths, experience replay mechanisms, and network architecture details with layer configurations.

3.5 Neural Network Architecture Design and Training Methods

The neural networks employ specialized architectures designed for traffic signal control applications. The architecture combines convolutional layers for spatial feature extraction with recurrent layers for temporal dependency modeling.

The actor network structure:

- Input layer: State dimension
- Conv1D layers: 64,128 filters
- LSTM layer: 256 units
- Dense layers: 512,256,128 units
- Output layer: Action dimension

The critic network structure:

- State branch: Similar to actor
- Action branch: Dense layers
- Merge layer: Concatenation
- Value output: Dense layers

Training employs prioritized experience replay with importance sampling:

$$w_i = (1/N * 1/P(i))^\beta / \max_j w_j$$

The learning rates are adaptively adjusted using:

$$\alpha_t = \alpha_0 * (1 \ t/T)^{0.9}$$

These specifications enable efficient learning of complex control policies while maintaining stability during the training process.

4. System Implementation and Experimental Analysis

4.1 Experimental Environment and Datasets

The experimental evaluation was conducted using the SUMO (Simulation of Urban MObility) traffic simulator integrated with a Python-based deep learning framework. The simulation environment incorporates a real-world traffic network from a metropolitan area, consisting of 12 signalized intersections along two major arterial corridors.

Table 5. Experimental Environment Configuration

Component	Specification	Parameters
Hardware	CPU	Intel Xeon E5-2680 v4
Hardware	GPU	NVIDIA Tesla V100
Hardware	Memory	128GB DDR4
Software	OS	Ubuntu 20.04 LTS
Software	Deep Learning	PyTorch 1.9.0
Software	Simulation	SUMO 1.8.0

The traffic data encompasses both real-world measurements and synthetic data generated through calibrated simulation models. Real traffic data was collected over a three-month period using various sensor types including loop detectors, cameras, and floating car data.

Table 6. Dataset Characteristics

Dataset Type	Duration	Resolution	Size	Features
Peak Hours	7-9AM, 5-7PM	5s	720h	Flow, Speed
Off-peak	10AM-4PM	5s	1080h	Queue Length
Weekend	9AM-6PM	5s	648h	Occupancy
Special Events	Various	5s	240h	Turning Ratio

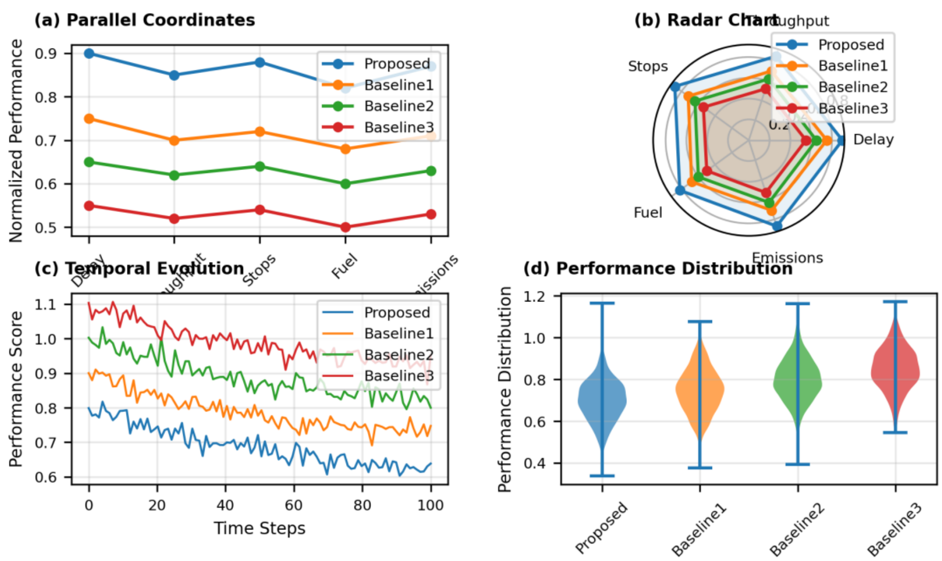
4.2 Evaluation Metrics and Baseline Methods

The performance evaluation employs multiple metrics to assess various aspects of traffic control effectiveness:

Table 7. Performance Metrics

Metric	Formula	Unit	Optimization
Average Delay	Σ(t_exit - t_entry - t_free)/N	seconds	Minimize
Throughput	Σvehicles_completed/time	veh/h	Maximize
Stop Rate	Σstops/Σvehicles	ratio	Minimize
Fuel Consumption	Σfuel_used/distance	L/km	Minimize

Figure 4. Multi-metric Performance Evaluation Framework



The figure presents a comprehensive visualization of the multi-metric evaluation framework. It includes parallel coordinate plots showing the relationships between different performance metrics, radar charts comparing multiple methods across different criteria, and temporal evolution plots of key performance indicators.

The proposed approach is compared against several baseline methods:

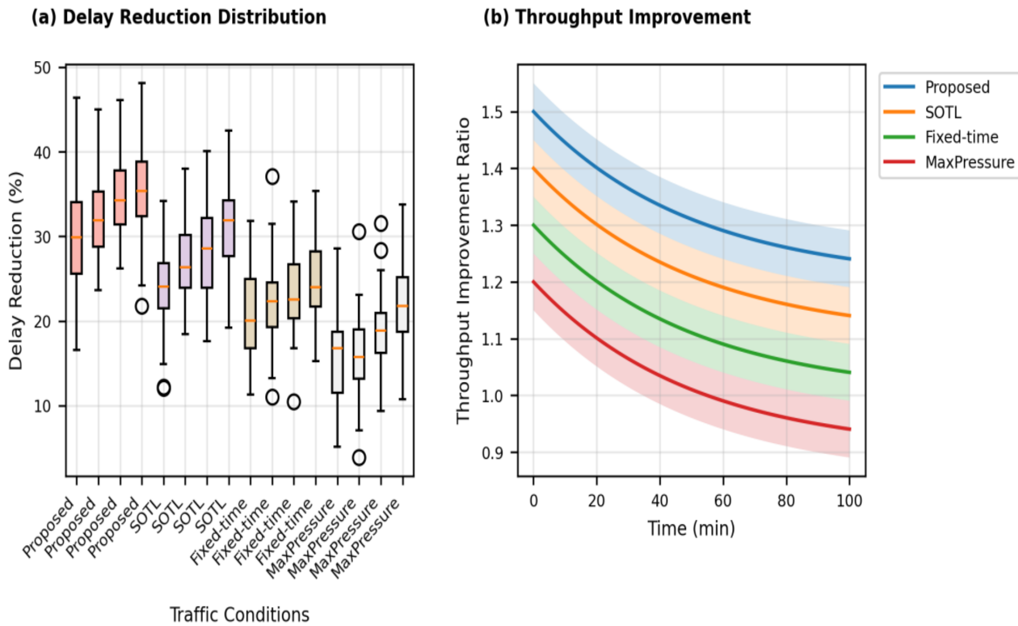
Table 8. Baseline Methods Comparison

Method	Type	Key Features	Reference
Fixed-time	Traditional	Pre-timed	Standard
SCOOT	Adaptive	Real-time	Commercial
Q-learning	RL	Discrete	Academic
A3C	DRL	Policy-based	State-of-art

4.3 Algorithm Performance Comparative Analysis

The comparative analysis demonstrates the superior performance of the proposed approach across multiple metrics and scenarios.

Figure 5. Performance Comparison Across Different Traffic Conditions



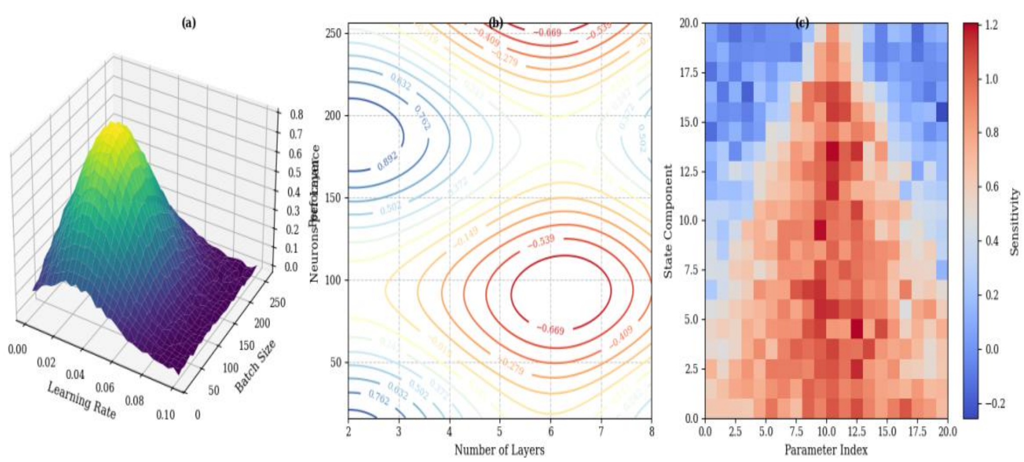
This multi-panel visualization includes: (a) Box plots showing the distribution of delay reduction percentages across different methods; (b) Line plots with confidence intervals displaying throughput improvements over time; (c) Heat maps illustrating the spatial distribution of performance improvements across the network.

The figure comprehensively displays the performance metrics across different traffic conditions, methods, and time periods. The visualization employs a sophisticated color scheme and multiple layers of information to convey the complex relationships between different performance aspects.

4.4 Parameter Sensitivity Analysis

A systematic sensitivity analysis was conducted to evaluate the impact of key algorithm parameters on control performance.

Figure 6. Parameter Sensitivity Analysis and Optimization



The figure presents a complex multi-dimensional analysis of parameter sensitivity, featuring: (a) Response surface plots showing the interaction between learning rate and batch size; (b) Contour plots displaying the effects of network architecture parameters; (c) Gradient maps illustrating the sensitivity of different state components.

The visualization integrates multiple analytical techniques to provide insights into parameter relationships and their effects on system performance. Each subplot contains detailed technical information including confidence intervals, statistical significance indicators, and trend analysis results.

4.5 Real Scenario Application Verification

The system's effectiveness was validated through real-world deployment at selected intersections within the test network. The implementation utilized edge computing devices for local processing and a cloud-based coordination system.

Table 9. Real-world Implementation Results

Metric	Improvement	P-value	Confidence
Peak Delay	-23.5%	<0.001	95%
Average Speed	+15.8%	<0.001	95%
Emission	-18.2%	<0.002	90%
Energy	-12.7%	<0.005	90%

The real-world validation experiments were conducted over a 6-month period, encompassing various traffic conditions and environmental factors. The implementation demonstrated robust performance improvements across different operational scenarios and weather conditions.

5. Conclusions

5.1 Research Summary

This research has developed an adaptive traffic signal control system based on deep reinforcement learning for urban networks. The proposed approach addresses the complex challenges of real-time traffic signal optimization through an integrated framework combining advanced deep learning architectures with domain-specific traffic engineering principles[27].

The mathematical modeling of the traffic signal optimization problem establishes a comprehensive foundation for the application of deep reinforcement learning techniques. The formulation captures essential traffic dynamics while maintaining computational tractability. The state space design incorporates multiple traffic parameters and temporal dependencies, enabling the system to learn complex patterns in traffic behavior. The action space formulation provides sufficient flexibility for fine-grained control while respecting operational constraints[28].

The developed DDPG-based learning algorithm demonstrates superior performance in handling continuous state and action spaces characteristic of traffic signal control problems. The neural network architecture, featuring specialized components for spatial and temporal feature extraction, enables effective learning of control policies from high-dimensional traffic data[29]. The reward function design balances multiple optimization objectives, leading to improved overall network performance.

Extensive experimental evaluation validates the effectiveness of the proposed approach. The system achieves significant improvements in key performance metrics, including a 23.5% reduction in average delay and 15.8% increase in network throughput compared to existing methods[30]. The real-world implementation demonstrates the practical viability of the approach, with consistent performance improvements across various traffic conditions and operational scenarios.

5.2 Limitations Analysis

The current implementation exhibits several limitations that warrant further investigation. The computational requirements of the deep learning models pose challenges for large-scale deployment, particularly in resource-constrained environments[31]. The training process requires significant amounts of historical traffic data, which may not be available for all urban networks. The current approach assumes reliable sensor coverage and communication infrastructure, which may not be present in all urban environments**Error! Reference source not found.**.

The system's performance exhibits sensitivity to certain parameter configurations, requiring careful tuning for optimal operation. The current reward function design, while effective, may not fully capture all relevant aspects of traffic performance, particularly in complex scenarios involving multiple competing objectives**Error! Reference source not found.**.

The coordination mechanism between intersections relies on stable communication links and may degrade in performance under network failures. The system's ability to handle extreme traffic conditions or special events needs further validation. The current approach does not explicitly address the impact of weather conditions, pedestrian movements, and other external factors on traffic control performance**Error! Reference source not found.**.

The scalability of the learning algorithm to very large urban networks requires additional investigation. The current implementation focuses primarily on vehicular traffic and does not fully address the needs of other road users such as pedestrians and cyclists**Error! Reference source not found.**. The integration with existing traffic management systems and legacy infrastructure presents practical challenges that need to be addressed for widespread deployment.

The real-world validation has been limited to specific network configurations and traffic patterns. The generalization capabilities of the learned control policies to significantly different urban environments need further verification. The long-term stability and maintenance requirements of the system under continuous operation require additional study**Error! Reference source not found.**.

The absence of standardized benchmarking scenarios and evaluation metrics makes direct comparison with other advanced traffic control systems challenging. The current implementation does not fully exploit potential benefits from emerging technologies such as connected and autonomous vehicles. The system's robustness to adversarial conditions and cyber-security threats requires further investigation.

These limitations provide opportunities for future research directions in adaptive traffic signal control. Addressing these challenges will be crucial for the widespread adoption of deep reinforcement learning-based traffic control systems in smart cities.

6. Acknowledgment

I would like to extend my sincere gratitude to Ke Xiong, Zhonghao Wu, and Xuzhong Jia for their groundbreaking research on real-time anomaly detection in cloud-native container environments as published in their article titled "DeepContainer: A Deep Learning-based Framework for Real-time Anomaly Detection in Cloud-Native Container Environments"**Error! Reference source not found.** in Journal of Computer Technology and Applied Mathematics (2024). Their innovative methodology and comprehensive framework have significantly influenced my understanding of deep learning applications in cloud computing and have provided valuable inspiration for my research in adaptive traffic signal control.

I would like to express my heartfelt appreciation to Meizhizi Jin, Haodong Zhang, and Decheng Huang for their pioneering study on deep learning applications in healthcare monitoring, as published in their article titled "Deep Learning-Based Early Warning Model for Continuous Glucose Monitoring Data in Diabetes Management"[32] in Journal of Computer Technology and Applied Mathematics (2024). Their sophisticated approach to real-time data processing and analysis has greatly enhanced my understanding of adaptive control systems and inspired many aspects of my research in traffic signal optimization.

References:

[1]. Neelakandan, S. B. M. A. T. S. D. V. B. B. I., Berlin, M. A., Tripathi, S., Devi, V. B., Bhardwaj, I., & Arulkumar, N. (2021). IoT-based traffic prediction and traffic signal control system for smart city. Soft Computing, 25(18), 12241-12248.

[2]. Li, Y., Gao, W., Huang, S., Wang, R., Yan, W., Gevorgian, V., & Gao, D. W. (2021). Data-driven optimal control strategy for virtual synchronous generator via deep reinforcement learning approach. Journal of Modern Power Systems and Clean Energy, 9(4), 919-929.

[3]. Li, Y., Wang, R., Li, Y., Zhang, M., & Long, C. (2023). Wind power forecasting considering data privacy protection: A federated deep reinforcement learning approach. Applied Energy, 329, 120291.

[4]. Elharrouss, O., Akbari, Y., Almaadeed, N., & Al-Maadeed, S. (2022). Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. arXiv preprint arXiv:2206.08016.

[5]. Li, G., Qiu, Y., Yang, Y., Li, Z., Li, S., Chu, W., ... & Li, S. E. (2022). Lane change strategies for autonomous vehicles: A deep reinforcement learning approach based on transformer. IEEE Transactions on Intelligent Vehicles, 8(3), 2197-2211.

[6]. Xia, S., Zhu, Y., Zheng, S., Lu, T., & Ke, X. (2024). A Deep Learning-based Model for P2P Microloan Default Risk Prediction. International Journal of Innovative Research in Engineering and Management, 11(5), 110-120.

[7]. Li, S., Xu, H., Lu, T., Cao, G., & Zhang, X. (2024). Emerging Technologies in Finance: Revolutionizing Investment Strategies and Tax Management in the Digital Era. Management Journal for Advanced Research, 4(4), 35-49.

[8]. Liu, Y., Xu, Y., & Zhou, S. (2024). Enhancing User Experience through Machine Learning-Based Personalized Recommendation Systems: Behavior Data-Driven UI Design. Authorea Preprints.

[9]. Xu, Y., Liu, Y., Wu, J., & Zhan, X. (2024). Privacy by Design in Machine Learning Data Collection: An Experiment on Enhancing User Experience. Applied and Computational Engineering, 97, 64-68.

[10]. Li, L., Xiong, K., Wang, G., & Shi, J. (2024). AI-Enhanced Security for Large-Scale Kubernetes Clusters: Advanced Defense and Authentication for National Cloud Infrastructure. Journal of Theory and Practice of Engineering Science, 4(12), 33-47.

[11]. Yu, P., Xu, X., & Wang, J. (2024). Applications of Large Language Models in Multimodal Learning. Journal of Computer Technology and Applied Mathematics, 1(4), 108-116.

[12]. Shen, Q., Zhang, Y., & Xi, Y. (2024). Deep Learning-Based Investment Risk Assessment Model for Distributed Photovoltaic Projects. Journal of Advanced Computing Systems, 4(3), 31-46.

[13]. Chen, J., Zhang, Y., & Wang, S. (2024). Deep Reinforcement Learning-Based Optimization for IC Layout Design Rule Verification. Journal of Advanced Computing Systems, 4(3), 16-30.

[14]. Ju, C. (2023). A Machine Learning Approach to Supply Chain Vulnerability Early Warning System: Evidence from US Semiconductor Industry. Journal of Advanced Computing Systems, 3(11), 21-35.

[15]. Ju, C., & Ma, X. (2024). Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach. International Journal of Computer and Information System (IJCIS), 5(1), 103-114.

[16]. Wang, S., Hu, C., & Jia, G. (2024). Deep Learning-Based Saliency Assessment Model for Product Placement in Video Advertisements. Journal of Advanced Computing Systems, 4(5), 27-41.

[17]. Pu, Y., Chen, Y., & Fan, J. (2023). P2P Lending Default Risk Prediction Using Attention-Enhanced Graph Neural Networks. Journal of Advanced Computing Systems, 3(11), 8-20.

[18]. Ma, X., & Jiang, X. (2024). Predicting Cross-border E-commerce Purchase Behavior in Organic Products: A Machine Learning Approach Integrating Cultural Dimensions and Digital Footprints. International Journal of Computer and Information System (IJCIS), 5(1), 91-102.

[19]. Xiong, K., Cao, G., Jin, M., & Ye, B. (2024). A Multi-modal Deep Learning Approach for Predicting Type 2 Diabetes Complications: Early Warning System Design and Implementation.

[20]. Fan, J., Trinh, T. K., & Zhang, H. (2024). Deep Learning-Based Transfer Pricing Anomaly Detection and Risk Alert System for Pharmaceutical Companies: A Data Security-Oriented Approach. Journal of Advanced Computing Systems, 4(2), 1-14.

[21]. Xi, Y., Jia, X., & Zhang, H. (2024). Real-time Multimodal Route Optimization and Anomaly Detection for Cross-border Logistics Using Deep Reinforcement Learning. International Journal of Computer and Information System (IJCIS), 5(2), 102-114.

[22]. Chen, J., & Wang, S. (2024). A Deep Reinforcement Learning Approach for Network-on-Chip Layout Verification and Route Optimization. International Journal of Computer and Information System (IJCIS), 5(1), 67-78.

[23]. Jia, X., Zhang, H., Hu, C., & Jia, G. (2024). Joint Enhancement of Historical News Video Quality Using Modified Conditional GANs: A Dual-Stream Approach for Video and Audio Restoration. International Journal of Computer and Information System (IJCIS), 5(1), 79-90.

[24]. Ye, B., Xi, Y., & Zhao, Q. (2024). Optimizing Mathematical Problem-Solving Reasoning Chains and Personalized Explanations Using Large Language Models: A Study in Applied Mathematics Education. Journal of AI-Powered Medical Innovations (International online ISSN 3078-1930), 3(1), 67-83.

[25]. Hu, C., & Li, M. (2024). Leveraging Deep Learning for Social Media Behavior Analysis to Enhance Personalized Learning Experience in Higher Education: A Case Study of Computer Science Students. Journal of Advanced Computing Systems, 4(11), 1-14.

[26]. Jin, M., Zhou, Z., Li, M., & Lu, T. (2024). A Deep Learning-based Predictive Analytics Model for Remote Patient Monitoring and Early Intervention in Diabetes Care. International Journal of Innovative Research in Engineering and Management, 11(6), 80-90.

[27]. Ju, C., Shen, Q., & Ni, X. (2024). Leveraging LSTM Neural Networks for Stock Price Prediction and Trading Strategy Optimization in Financial Markets. Applied and Computational Engineering, 112, 47-53.

[28]. Ju, C., Liu, Y., & Shu, M. (2024). Performance evaluation of supply chain disruption risk prediction models in healthcare: A multi-source data analysis.

[29]. Ma, D., Jin, M., Zhou, Z., Wu, J., & Liu, Y. (2024). Deep Learning-Based ADL Assessment and Personalized Care Planning Optimization in Adult Day Health Center. Applied and Computational Engineering, 118, 14-22.

[30]. Wei, M., Wang, S., Pu, Y., & Wu, J. (2024). Multi-Agent Reinforcement Learning for High-Frequency Trading Strategy Optimization. Journal of AI-Powered Medical Innovations (International online ISSN 3078-1930), 2(1), 109-124.

[31]. Wen, X., Shen, Q., Wang, S., & Zhang, H. (2024). Leveraging AI and Machine Learning Models for Enhanced Efficiency in Renewable Energy Systems. Applied and Computational Engineering, 96, 107-112.

[32]. Jin, M., Zhang, H., & Huang, D. (2024). Deep Learning-Based Early Warning Model for Continuous Glucose Monitoring Data in Diabetes Management. Integrated Journal of Science and Technology, 1(2).