Sensors & Transducers, Vol. 270, Issue 3, November 2025, pp. 60-68

Sensors & Transducers
Published by IFSA Publishing, S. L., 2025
http://www.sensorsportal.com

Integrating Visible Light Communication and Deep Reinforcement Learning for Smarter Urban Traffic Control

[1] Gonçalo GALVÃO, [1, 2] Manuel A. VIEIRA, [1, 2, 3] Manuela VIEIRA, [1, 4] Mário VÉSTIAS, [1, 2] Paula LOURO and [1, 5] Pedro VIEIRA
[1] Electronics Telecommunication and Computer Dept. ISEL, R. Conselheiro Emídio Navarro, 1959-007 Lisboa, Portugal
[2] UNINOVA –CTS and LASI, Quinta da Torre, Monte da Caparica, 2829-516, Caparica, Portugal
[3] NOVA School of Science and Technology, Quinta da Torre, Monte da Caparica, 2829-516, Caparica, Portugal
[4] INESC INOV-Lab, Instituto Superior Técnico, Universidade de Lisboa, 1000-029, Lisboa, Portugal
[5] Instituto de Telecomunicações, Instituto Superior Técnico, 1049-001, Lisboa, Portugal
E-mail: manuela.vieira@isel.pt

Received: 26 May 2025 / Revised: 6 Nov. 2025 / Accepted: 12 Nov. 2025 / Published: 28 Nov. 2025

Abstract: Urban traffic management is increasingly challenged by rising vehicle and pedestrian flows, resulting in congestion, delays, and safety riskW. This article proposes an innovative traffic signal control framework that integrates Deep Reinforcement Learning (DRL) with Visible Light Communication (VLC) to optimize operations at intersections, which are critical bottlenecks in urban networks. A decentralized DRL agent is deployed at each intersection and trained on local traffic states, enabling real-time decision-making for both vehicular and pedestrian movements. VLC is used to support low-latency, infrastructure-to-user communication, providing accurate data on positions, speeds, queue lengths, and stop durations. The system employs a queue/request/response mechanism to adapt signal phases dynamically, resolve conflicts, and prioritize urgent flows. The proposed approach is validated through simulations and real-world trials, demonstrating superior performance over centralized and traditional agent-based methods by substantially reducing waiting and travel times while enhancing safety. The solution is scalable and adaptable to a wide range of intersection configurations, with SUMO-based experiments confirming its potential for more efficient and intelligent urban traffic control.

Keywords: Urban traffic management, Deep reinforcement learning (DRL), Visible light communication (VLC), Decentralized traffic control, Real-time decision making, Vehicle/pedestrian-to-Infrastructure communication (V/P2I), Road safety.

1. Introduction

Every day, people traveling within cities face considerable delays in reaching their destinations as a direct consequence of traffic congestion. The continuous and accelerated growth of urban populations has inevitably resulted in a substantial increase in the number of vehicles circulating within city environments. This growth exerts additional pressure on existing road infrastructures, leading to longer queues, increased travel times, and a general decline in transportation efficiency. The resulting congestion not only affects individual mobility but also has broader socioeconomic and environmental implications, such as elevated fuel consumption, increased greenhouse gas emissions, and reduced productivity.

Despite these challenges, the evolution of traffic management systems has not kept pace with the rapid changes in mobility demand. The expansion of vehicle fleets and the growing complexity of urban transport networks have not been matched by proportional investment in, or adoption of, intelligent and adaptive traffic control technologies. In many metropolitan areas, including cities such as Lisbon, the traffic management infrastructure remains largely outdated and rigid, relying predominantly on pre-timed, cyclic phase control strategies that operate under fixed schedules regardless of real-time fluctuations in traffic flow.

Urban traffic management is a critical challenge for modern cities, particularly during peak entry and exit hours when vehicle and pedestrian volumes significantly increase, leading to severe congestion, delays, and safety risks. Since expanding road infrastructure is often impractical, intelligent optimization of traffic flow at intersections – especially during these critical time windows – has become essential. One of the most promising strategies is the implementation of adaptive traffic signal control systems that adjust signal timings in real time based on current traffic conditions such as vehicle queues and pedestrian flows.

Deep Reinforcement Learning (DRL) has emerged as an effective approach for dynamic signal control, capable of learning optimal policies through continuous interaction with the environment. However, managing multiple intersections in urban networks remains a complex task due to heterogeneous traffic patterns and the need for coordination across nodes [1, 2].

In this context, Visible Light Communication (VLC) presents an innovative solution to support real-time data exchange using the intensity modulation of Light Emitting Diodes (LEDs), already integrated into streetlights, traffic signals, and vehicle lighting systems. This dual-functionality allows existing infrastructure to serve both lighting and communication purposes, enabling seamless interaction between road infrastructure and vehicles or pedestrians.

By integrating DRL with VLC and Connected Vehicle (CV) technologies, it becomes possible to achieve coordinated, real-time traffic management across intersections. The research question guiding this study is:

How the combined use of DRL, VLC, and CVs can enhance the responsiveness and efficiency of urban traffic systems, especially in adapting to dynamic demands during daily rush hours.

This research investigates the integration of VLC and AI, with a focus on DRL-based strategies, for adaptive traffic control. Specifically, we explore the design and training of neural networks capable of determining dynamic phase timings using the SAPA module developed in this study. The system applies DRL principles within a connected vehicular ecosystem to improve intersection performance, optimize arterial traffic flow, and balance the operational needs of both vehicles and pedestrians in real-world urban scenarios. By leveraging VLC and CVs, this work aims to demonstrate the potential of emerging communication and learning technologies to address modern urban traffic management challenges effectively.

The remainder of this article is organized as follows. Section 2, reviews the related work, including the impact of VLC on Intelligent Transportation Systems, recent advances in Deep Reinforcement Learning for traffic control, and existing traffic control strategies. Section 3, presents the proposed traffic signal control framework for multi-intersection networks, describing the experimental environment, the considered traffic control strategies, and the system architecture together with the VLC communication model. Section 4, discusses the results obtained from the experimental evaluation, highlighting the performance and implications of the proposed approach. Finally, Section 5, concludes the article by summarizing the main contributions and outlining future research directions.

2. Related Work

2.1. VLC impact on Intelligent Transportation Systems

Visible Light Communication (VLC) is gaining traction as a complementary technology in the evolution of Intelligent Transport Systems (ITS), supporting a range of applications such as communication at signalized intersections, collision warning and avoidance, vehicle localization, and platooning through V2V, I2V, and broader V2X communications [3, 4].

While RF communication plays a foundational role in ITS, it faces notable limitations in dense urban environments. Issues such as electromagnetic interference from electronic devices, spectrum congestion due to high demand, and security vulnerabilities can compromise RF performance. Its openness also makes it susceptible to eavesdropping and cyberattacks.

VLC, by contrast, uses modulated LED light to transmit data, providing several benefits. It is immune to electromagnetic interference, operates in an unlicensed and uncongested spectrum, and offers enhanced data security. Additionally, VLC systems can be implemented using existing lighting infrastructure, making them a cost-effective alternative to RF in many contexts [5, 6].

While VLC provides substantial advantages for short-range, high-capacity communication in dense urban environments, it also presents intrinsic challenges. Because data transmission relies on light modulation, simultaneous emissions from multiple transmitters (e.g., vehicles and infrastructure units) can cause interference and data collisions, especially in heavy-traffic scenarios. Such interference can degrade signal quality and reliability, potentially compromising real-time vehicular communications. To mitigate these effects, several techniques have been proposed, including Optical Code Division Multiple Access (OCDMA) to reduce multi-user interference [7], advanced modulation and adaptive power control schemes to enhance link stability, and hybrid VLC/RF architectures that dynamically switch between optical and radio channels depending on network and environmental conditions [8-10].

However, VLC is not without its own challenges – particularly in outdoor scenarios. Environmental conditions such as bright sunlight, fog, rain, or snow can degrade signal quality. One of the most critical issues is the saturation of optical receivers caused by direct sunlight, which significantly reduces system performance.

Given these limitations, VLC should not be viewed as a replacement for RF communication but rather as a complementary technology. Combining the strengths of both approaches can lead to more robust, secure, and reliable ITS deployments.

2.2. Deep Reinforcement Learning in Traffic Control Systems

Managing traffic flow efficiently is crucial to keeping cities running smoothly. Without proper strategies in place, urban road networks can become overwhelmed, leading to significant delays and longer journeys. Recent progress in artificial intelligence has brought Deep Reinforcement Learning (DRL) to the forefront as a promising solution for traffic control in vehicular communication systems. DRL leverages deep neural networks to analyze complex traffic behaviors and adapt signal timings dynamically, enabling smarter and more responsive control mechanisms [11].

The observation and decision-making processes in such systems are carried out by agents, with the number of agents depending on the complexity of the traffic scenario. Simpler environments, such as those with a single intersection, often rely on a single DRL agent. In contrast, larger and more intricate networks involving multiple intersections typically require a Multi-Agent Reinforcement Learning (MARL) framework. In these setups, it becomes essential to consider how agents interact and make decisions – whether the control is centralized or distributed, whether the agents have full or limited visibility of the environment, and whether they collaborate or compete.

In a centralized setup, a single controller dictates the actions for all agents at every decision point. Decentralized models, on the other hand, assign autonomy to each agent, allowing them to act independently. Agents may work together towards a shared objective or operate individually to pursue their own rewards. Their knowledge of the environment might be restricted to local observations or supplemented by information exchanged with neighboring agents.

The environment in reinforcement learning is commonly represented as a Markov Decision Process (MDP), defined by the tuple <S, A, P, R, γ>. In this model, *S* is the set of all possible states, *A* the available actions, *P* the probabilities of transitioning between states, *R* the function that assigns rewards, and γ the discount factor that balances the importance of immediate versus future rewards. At each time step t, an agent perceives its current state s_t, performs an action a_t, receives a reward r_t, and transitions to the next state s_{t+1}. Over time, the agent refines its strategy to maximize the cumulative reward, reinforcing behaviors that yield better outcomes.

2.3. Traffic Control Strategies

Traffic control strategies can be categorized based on how they operate and the context in which they are applied [11]. A key distinction lies between fixed-time and traffic-responsive approaches. Fixed-time strategies rely on offline optimization using historical traffic data for specific periods (e.g., morning rush hours), assuming constant demand and turning rates. In contrast, traffic-responsive systems use real-time traffic data – typically from inductive loop detectors – to adjust signal timings dynamically.

Another distinction is between isolated and coordinated control. Isolated strategies focus on individual intersections, while coordinated systems manage groups of intersections across urban zones or entire networks to improve overall flow.

Most traditional control strategies are designed for undersaturated conditions, where traffic queues form during red lights but clear during green phases. However, few methods are equipped to handle oversaturated situations, where queues persist and can even extend upstream to other intersections.

Within fixed-time strategies, stage-based methods optimize signal splits and cycle times to reduce delays or increase capacity, while phase-based strategies go further by also optimizing the signal staging – particularly useful at complex intersections.

Despite their widespread use, fixed-time systems have notable limitations. Since they depend on historical data, they often fail to reflect real-time conditions. Traffic demand can fluctuate throughout the day, differ across days due to events, evolve over time, or shift in response to changes in signal settings. Moreover, unexpected incidents or disruptions can significantly alter traffic patterns in ways that static strategies cannot accommodate.

3. Traffic Signal Control at Multi-intersection Networks

Arterial traffic signal control refers to managing intersections formed by crossing two or more main roads, either radial or circular in design. The layout and spacing between intersections vary depending on traffic volume, road capacity, and network design. Each approach at an intersection comprises multiple lanes to accommodate different vehicle movements such as left-turns, right-turns, and through-traffic. These intersections are governed by standard traffic rules, with priority movements determined by the traffic signals in place.

3.1. Experimental Environment

We consider a simplified urban traffic network composed of two intersecting arterial roads: a horizontal road (C0-C1-C2) and a vertical road (C3-C1-C4), which meet at the central junction C1 [6]. Intersection C1 serves as the unique connection point between these "horizontal" and "vertical" arteries. In the setup (Fig. 1), C1 has no local sources of traffic – it only receives vehicles from the four adjacent intersections. All incoming flows into C1's lanes are determined by the phase-activation decisions of the neighboring intersection controllers (agents C0, C2, C3, and C4) In effect, C1's role is to mediate the streams from the two arteries by influencing how and when those neighboring agents release traffic. This configuration makes C1 a central hub whose activity can substantially affect the overall network dynamics.

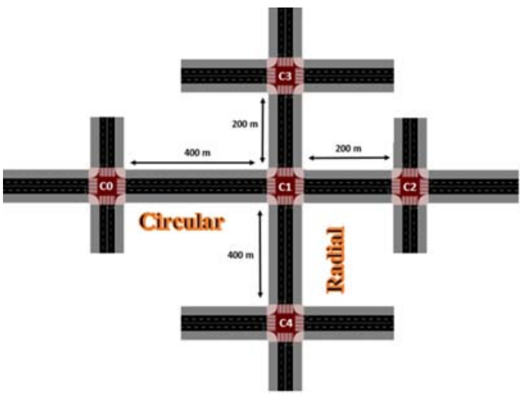


Fig. 1. Traffic scenario consisting of 5 homogeneous intersections with 4 arms each.

In this case, C1's coordination might promote balanced traffic dispersal or alleviate congestion, but it could also introduce imbalances if misaligned. We therefore investigate how different priority schemes imposed at C1 affect traffic flow through the whole network.

This configuration of five intersections is designated as a cell within the traffic management system.

3.2. Traffic Control Strategies

Five distinct control strategies are implemented, each via a separate neural network (agent) that biases the traffic phases at the adjacent intersections. These strategies differ in how they prioritize the two arteries (horizontal vs. vertical) and the direction of flow (inbound vs. outbound relative to C1). The goal is to compare a balanced approach with schemes that favor one road or direction. Specifically:

- Strategy 1 – Balanced Strategy: No prioritization. Both arteries and directions are treated equally to maximize overall fairness and throughput;
- Strategy 2 – Circular + Outbound Radial: Prioritizes the horizontal (circular) road and outbound flow on the vertical axis (S→N), favoring east–west and northbound movements;
- Strategy 3 – Circular + Inbound Radial: Emphasizes the circular road and inbound flow on the vertical axis (N→S), giving priority to vehicles entering the center from the north and to east–west movement;
- Strategy 4 – Radial + Outbound Traffic: Focuses on the vertical (radial) road and outbound traffic (northbound), promoting north–south flow and reducing congestion along the vertical axis;
- Strategy 5 – Radial + Inbound Traffic: Gives priority to inbound traffic along the radial road (southbound) and to east–west movements, supporting flow toward the center.

Each of these strategies embodies a different hypothesis about traffic demand: some give circular road an advantage, others the radial road, and each distinguishes whether to push traffic out of the center (outbound) or pull it into the center (inbound). The implementation of these biases is achieved by modifying the generation of the vehicles in the environment. In the case of prioritizing a direction, like for example W-E movements, then more vehicles are being generated in C0 and C2, West and East roads, respectively. The same procedure is also done for the radial artery. With this each one of the trained networks is adapted to a specific strategy as summarized in Table 1. For example, and considering a total traffic demand of 1800 vehicles, in all strategies under consideration, 75 % of the vehicles (1350 vehicles) proceed straight ahead or make right turns, while the remaining 25 % (450 vehicles) correspond to left-turn movements.

In Strategies 2 and 3 of the 1350 vehicles that proceed straight or turn right, 65 % (878 vehicles) are generated on the circular artery, whereas the remaining 35 % (472 vehicles) originate from the radial artery. To represent inbound and outbound city movements, traffic generation on the radial artery is deliberately unbalanced: 75 % of these 472 vehicles (354 vehicles) are produced in the south–north direction (outbound flow) at intersection C4, with the remaining 25 % (118 vehicles) generated at intersection C3, or conversely in the case of inbound flows.

Similarly, in Strategies 4 and 5, the same distribution logic applies, but with 65 % of the 1350 vehicles (878 vehicles) generated on the radial artery. Again, this generation is unbalanced to simulate inbound and outbound flows, such that 75 % (659 vehicles) are generated at C4 and 25 % (219 vehicles) at C3. The circular artery, in turn, accommodates the remaining 35 % of the 1350 vehicles (472 vehicles).

Table 1. Summary of traffic control strategies applied.

	Strategy	Priority Artery	Direction Focus	Description
1	Standard	None	W-E N-S	Arteries and directions treated equally.
2	Circular + Outbound Radial	Circular	E-W Northbound (C4->C3)	Prioritizes circular artery with outbound radial flow (S→N).
3	Circular + Inbound Radial	Circular	E-W Southbound (C3->C4)	Prioritizes circular artery with inbound radial flow (N→S).
4	Radial + Outbound Radial	Radial	N–S Northbound (C4->C3)	Prioritizes radial artery with outbound radial flow.
5	Radial + Inbound Radial	Radial	N–S Southbound (C4->C3)	Prioritizes radial artery with inbound radial flow.

The experimental setup considered traffic flows of 1800 vehicles and 2000 pedestrians per hour, simulated across 200 episodes of 3600 seconds each. We will analyze how these five strategies impact learning and performance by comparing training and testing outcomes. In particular, we will measure the final cumulative reward achieved by each network and indicators of traffic fluidity (such as average vehicle halting time or throughput). This analysis will reveal which prioritization schemes help or hinder smooth circulation, illuminating whether intersection C1 can effectively act as a global coordinator in this arterial traffic scenario.

3.3. System Architecture and VLC Communication Model

The VLC system features modulated LED transmitters and multilayers photo receivers detecting light and wavelength variations. Light is modulated via ON-OFF Keying (OOK) [8]. Fig. 2 depicts the various VLC-based communication links considered in the system. It supports Infrastructure-to-Cloud (I2IM) and Vehicle-to-Vehicle (V2V) communication via embedded computing.

Streetlights, spaced 20 m apart, act as geo-transmitters, broadcasting L2V messages with IDs, synchronization, and traffic data. Vehicles/pedestrians receive a unique ID.

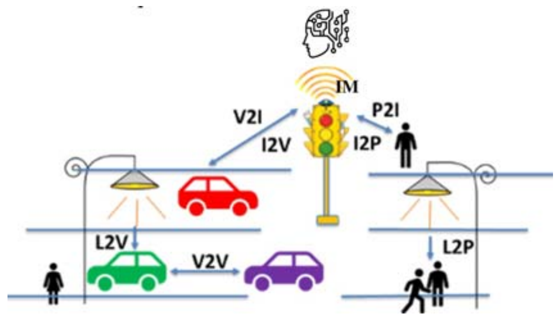


Fig. 2. Overview of visible light communication channels within the intersection area.

Intersection access follows a request/response protocol: a vehicle/pedestrian (V/P2I) sends a crossing request, and the infrastructure (I2V/P) replies with a trajectory assignment within footprint regions. If collision risk is detected, responses are delayed until safe. The response message informs the vehicle or pedestrian whether it is safe to proceed through the intersection, and it is issued by the Intersection Manager (IM).

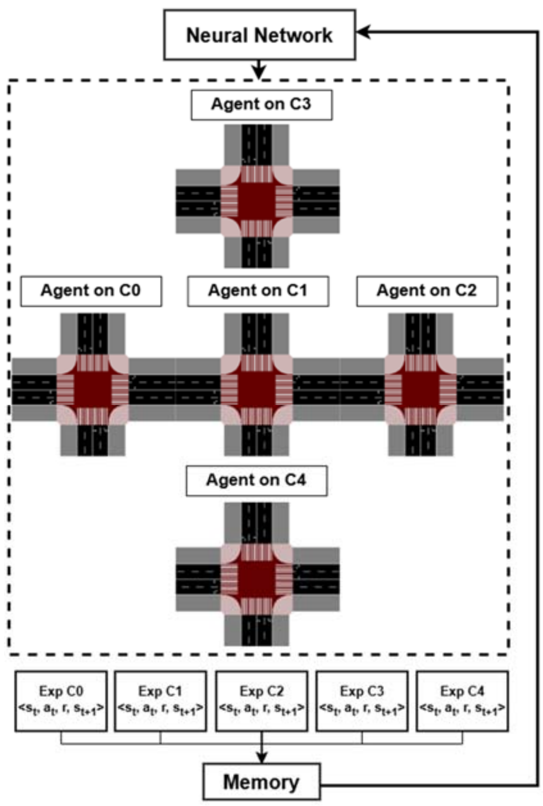


Fig. 3. Intersection Manager architecture based on a Deep Neural Network.

Fig. 3 illustrates the architecture of the IM, which is composed of a decentralized neural network trained based on the observations and experiences of individual agents. Each agent is responsible for controlling its own intersection. This neural network enables real-time decision-making, dynamically adjusting the active signal phases according to the observed traffic flows on each approach, thereby optimizing traffic movement within the cell.

Each agent performs local observations of its corresponding intersection and makes decisions regarding which signal phases to activate based on the perceived traffic state. The experiences collected by the agents are stored in a centralized replay memory to support the training of the neural network. This neural network, responsible for controlling the cell, is trained under a specific traffic control strategy, allowing it to become effectively adapted to the traffic dynamics characteristic of that strategy. Considering five distinct strategies, a dedicated neural network is trained for each one, resulting in five fully adapted models. These models are subsequently compared to evaluate and analyze their behavioral and performance differences across varying traffic scenarios.

4. Results and Discussion

The VLC system ensures real-time monitoring of pedestrians, vehicles, and infrastructure. Key metrics, such as queue formation and pedestrian density at corners, are evaluated to enhance safety. P2I2P communication facilitates travel time estimation, while transmitter tracking IDs provide insights into speed and waiting times [13, 14].

Fig. 4 demonstrates the MUX signal and the decoded messages between the vehicles and the traffic lights, respectively.

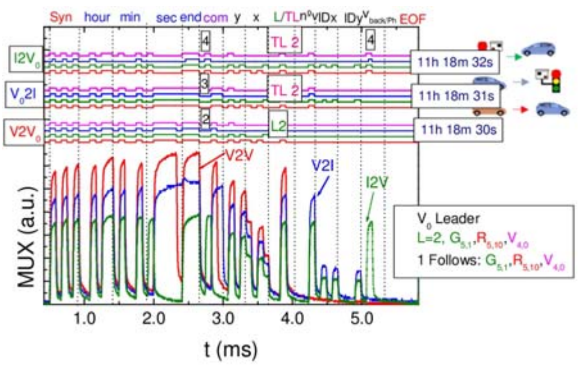


Fig. 4. MUX signal and the decoded messages.

Results show that with VLC is possible to details the flow of V2I, V2V, P2I, and I2P communications at various intersections, illustrating a structured communication framework for coordinating traffic and pedestrian movement.

For the traffic simulations, the implemented neural network follows a Multilayer Perceptron (MLP) architecture, consisting of an input layer with 164 neurons, 2 hidden layers with 400 neurons each, and an output layer with 9 neurons corresponding to the possible actions. All layers are Fully Connected (FCL), and the ReLU activation function is applied in the hidden layers. The network is trained using a Deep Q-Learning (DQL) algorithm, allowing it to learn optimal action policies based on observed state-action-reward transitions. Additional details regarding the network architecture and training configuration are provided in Table 2.

Table 2. Parameters of the neural network architecture and training configuration.

Parameter	Value
Episodes	200
Max Steps	3600
Vehicles	1800
Pedestrians	2000
Width Layers	400
Batch Size	100
Learning Rate	0.001
Loss Function	MSE
Training Epochs	500

In training, each of the networks learns to interact with the environment, each dealing with a specific traffic generation, corresponding to the strategy defined.

Fig. 5 presents the C1 training rewards of the 5 strategies and the respective testing metrics. With respect to the cumulative negative rewards, it is evident that each of the trained neural networks successfully converged and adapted to its respective strategy, effectively optimizing the traffic flow within the simulated environment. During the network evaluations, traffic metrics such as halted vehicles and pedestrian flow showed promising results. Each network learned a specific control strategy and, when tested, demonstrated the ability to effectively manage traffic, even during periods of high congestion. The models are capable of selecting the most suitable actions based on real-time traffic conditions at the intersection, thereby optimizing the flow of both vehicles and pedestrians.

For both vehicle and pedestrian metrics, it is observed that the strategy maintaining a balanced 50/50 allocation between the two main arteries results in the highest number of vehicles and pedestrians waiting. Since this strategy does not prioritize any specific direction, it activates signal phases in a uniformly distributed manner, which leads to increased vehicle accumulation – particularly at the critical intersection, C1. However, this does not imply that the strategy is ineffective; it serves a specific function within the overall structure of the traffic cell network.

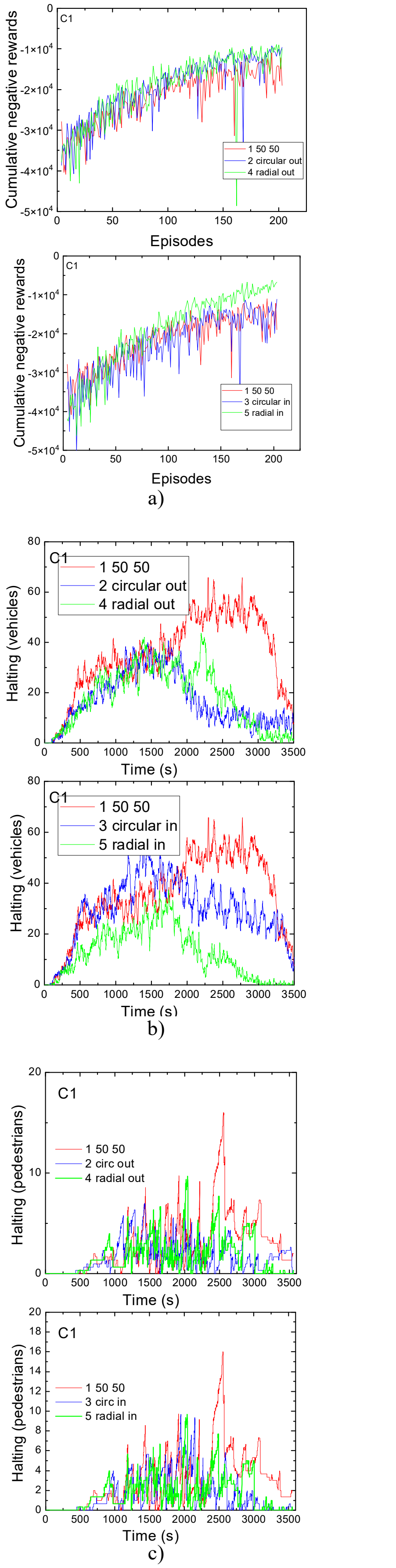


Fig. 5. C1 training rewards (a) and the respective testing vehicle (b) and Pedestrians (c) metrics of the five strategies.

In scenarios where a cell is located farther from the city center, it may be more appropriate to adopt strategy 3, which prioritizes circular traffic flow. As traffic approaches the urban core, control must tighten to prevent congestion. A gradual transition through strategy 1 and then strategy 5 allows for progressively increasing priority on radial flows, ensuring safer and smoother city entry.

The same logic applies for outbound traffic. Starting at the urban core, strategy 4 prioritizes radial flow (e.g., south-to-north). A shift to strategy 1 in the next cell introduces more balance, and finally strategy 2 – further from the center – emphasizes circular flow again. This strategic progression enables efficient redistribution of traffic, reducing congestion and improving overall flow.

Regarding the nine possible traffic phases, illustrated in Fig. 6.

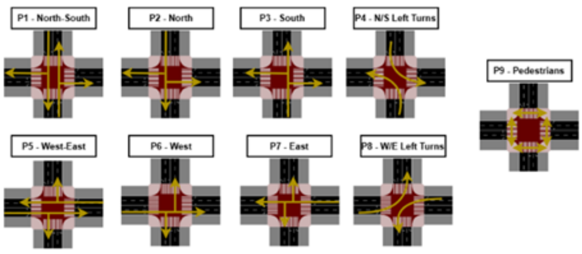
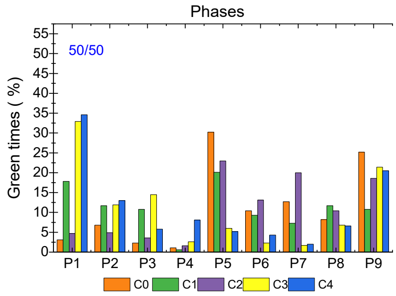


Fig. 6. Diagram of the nine possible traffic phases considered.

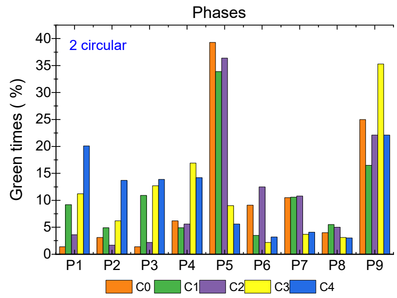
Of the nine total signal phases, eight are allocated to vehicular movements, while a single phase is reserved exclusively for pedestrian crossings. During this pedestrian phase, all vehicular movements are suspended, enabling pedestrians to traverse designated crosswalks safely without conflicts with vehicles. Pedestrian crossings are strictly restricted to their corresponding phase. This configuration not only optimizes traffic flow but also substantially enhances safety for both pedestrians and vehicles.

Fig. 7 presents the percentage distribution of the most frequently activated phases over time for each network. Strategy 1 demonstrates a balanced activation of the main signal phases, namely P1 (N–S direction), P5 (W–E direction), and P9 (pedestrian-exclusive phase), indicating an even distribution of traffic flow across the intersection approaches. In contrast, Strategy 2 handles a higher vehicle inflow along the W–E direction, prompting the system to favor phase P5 more frequently – especially at intersections C0, C1, and C2 – where its activation reaches an average of 35 %. Phase P1, associated with the radial N–S direction, sees reduced activation levels, typically between 10 % and 20 %, reflecting the shift in traffic demand. Phase P9 varies between 18 % and 35 %, depending on the intersection, though C1 consistently records the lowest pedestrian phase usage (below 20 %) across all networks. This behavior is expected, as C1 is a critical junction requiring continuous vehicular flow to avoid queue buildup and prevent spillback that could affect adjacent intersections.

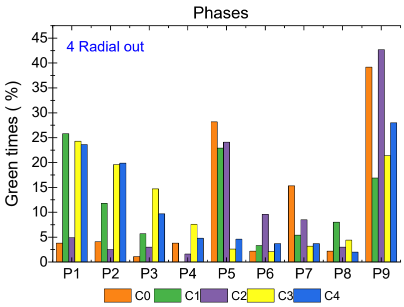
Strategy 4, characterized by higher traffic volumes along the radial axis, leads to increased activation of phase P1 (N–S), reaching around 25 % at radial intersections. In contrast, phase P5 decreases to similar levels (~25 %), while P9 varies between 18 % and 43 %, depending on the intersection. Although P1 and P5 exhibit comparable activation rates, resembling the balance seen in strategy 1, the dynamics differ. Strategy 4 also relies heavily on intermediate phases – P2 and P3 – to sustain radial flow, with activations ranging from 13–20 % and 6–15 %, respectively.



a)



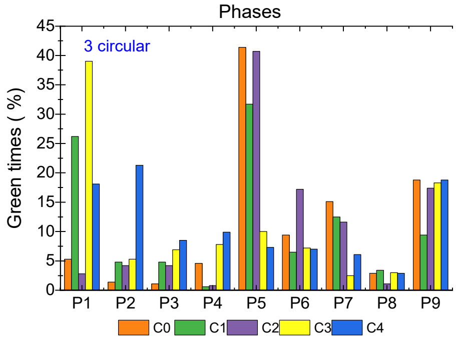
b)



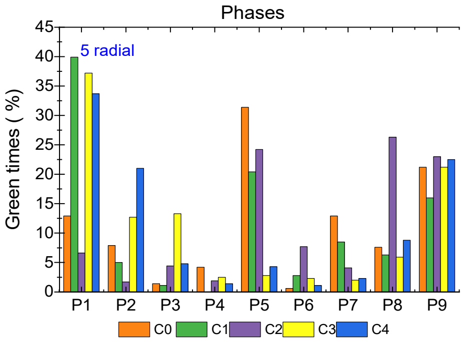
c)

Fig. 7. Comparison of active phases for each network under strategies a) 1 50/50, b) 2 Circular Out, c) 4 Radial Out.

Fig. 8 presents the strategies marked by strong inbound traffic into the city. In strategy 3, which prioritizes circular flow with entry via the radial artery, there is a significant increase in the activation of phase P5 (W–E), ranging from 30 % to 40 %. Phase P1 (N–S) also sees regular activation, between 18 % and 39 %, with intersection C3 showing the highest P1 activation – expected due to its role as a key traffic generator on the radial route.



a)



b)

Fig. 8. Comparison of active phases for each network under strategies a) 3 Circular In, b) 5 Radial In.

In strategy 5, which explicitly prioritizes the radial artery for inbound movement, phase P1 becomes dominant, with activation rates between 33 % and 39 %, surpassing those in strategy 3. This reflects the system's adaptation to traffic needs, ensuring efficient vehicle entry and sustained N–S flow. This also helps reduce queuing at C1, where increased N–S phase activation alleviates local and downstream congestion. Meanwhile, phase P5 decreases to 20–30 %, consistent with its reduced priority in this configuration.

5. Conclusions

This work presents a traffic signal control framework that integrates Deep Reinforcement Learning (DRL) with Visible Light Communication (VLC) to support real-time, adaptive decision-making in urban environments. By deploying decentralized DRL agents trained on local traffic data and enhanced by VLC-based communication, the system effectively manages vehicle and pedestrian flows under dynamic conditions. The evaluation of multiple network configurations, each trained under a specific strategy, demonstrates consistent improvements across both vehicle- and pedestrian-related metrics, including significant reductions in waiting times and queue lengths. Building on these results, future work will focus on the design, planning, and coordinated management of interconnected traffic cells, enabling scalable, city-wide optimization and more resilient multimodal traffic control.

Acknowledgments

This work was sponsored by FCT within the Research Unit Center of Technology and Systems CTS/UNINOVA/FCT/NOVA CTS/00066 and IPL/IDI&CA2024/INUTRAM_ISEL.

References

[1]. H. Ge, Y. Song, C. Wu, J. Ren, et al., Cooperative deep Q-learning with Q-value transfer for multi-intersection signal control, *IEEE Access*, Vol. 7, 2019, pp. 40797-40809.

[2]. A. Oroojlooy, D. Hajinezhad, A review of cooperative multi-agent deep reinforcement learning, *Applied Intelligence*, Vol. 53, 2023, pp. 15838-15883.

[3]. P. H. Pathak, X. Feng, P. Hu, P. Mohapatra, Visible light communication, networking, and sensing: potential and challenges, *IEEE Communications Surveys & Tutorials*, Vol. 17, Issue 4, 2015, pp. 2047-2077.

[4]. G. Galvão, M. A. Vieira, M. Vieira, M. Véstias, et al., Smart traffic signal control in urban networks via deep reinforcement learning and visible light communication, in *Proceedings of the 7th International Conference on Microelectronic Devices and Technologies (MicDat'25)*, 2025, pp. 28-34.

[5]. P. Sikder, M. T. Rahman, A. S. M. Bakibillah, Advancements and challenges of visible light communication in intelligent transportation systems: a comprehensive review, *Photonics*, Vol. 12, 2025, 225.

[6]. S. Caputo, et al., Measurement-based VLC channel characterization for I2V communications in a real urban scenario, *Vehicular Communications*, Vol. 28, 2021, 100305.

[7]. C. Zhu, M. Dastani, S. Wang, A survey of multi-agent deep reinforcement learning with communication, *Autonomous Agents and Multi-Agent Systems*, Vol. 38, 2024, 4.

[8]. E. Plascencia, H. Guan, L. Chassagne, A. M. Căilean, et al., Addressing multi-user interference in vehicular visible light communications: a brief survey and an evaluation of optical CDMA MAC utilization in a multi-lane scenario, *Sensors*, Vol. 23, Issue 8, 2023, 3831.

[9]. L. Bravo Alvarez, S. Montejo-Sánchez, L. Rodríguez-López, C. Azurdia-Meza, et al., A review of hybrid VLC/RF networks: features, applications, and future directions, *Sensors*, Vol. 23, Issue 17, 2023, 7545.

[10]. M. Garai, M. Sliti, N. Boudriga, M. Mrabet, et al., Optimized vehicular connectivity and data exchange in a tree-structured VLC communication network based on optical codewords, *Frontiers in Physics*, Vol. 13, 2025, 1635345.

[11]. M. Papageorgiou, Overview of road traffic control strategies, *IFAC Proceedings Volumes*, Vol. 37, Issue 19, 2004, pp. 29-40.

[12]. M. A. Vieira, M. Vieira, P. Vieira, P. Louro, Optical signal processing for a smart vehicle lighting system using a-SiCH technology, *Proceedings of SPIE*, Vol. 10231, 2017, 102311L.

[13]. M. A. Vieira, M. Vieira, P. Louro, P. Vieira, et al., Vehicular visible light communication for intersection management, *Signals*, Vol. 4, Issue 2, 2023, 457.

[14]. M. Vieira, M. A. Vieira, G. Galvão, P. Vieira, et al., Enhancing urban intersection efficiency: utilizing visible light communication and learning-driven control for improved traffic signal performance, *Vehicles*, Vol. 6, 2024, pp. 666-692.