vehicles

MDPI

Article

Integration of Visible Light Communication, Artificial Intelligence, and Rerouting Strategies for Enhanced Urban Traffic Management

Manuela Vieira [1,2,3,*], Gonçalo Galvão [1], Manuel A. Vieira [1,2], Mário Véstias [1,4], Pedro Vieira [1,5] and Paula Louro [1,2]

[1] DEETC-ISEL/IPL, R. Conselheiro Emídio Navarro, 1949-014 Lisboa, Portugal; a45903@alunos.isel.pt (G.G.); mv@isel.pt (M.A.V.); mario.vestias@isel.pt (M.V.); pedro.vieira@isel.pt (P.V.); paula.louro@isel.pt (P.L.)
[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
* Correspondence: mv@isel.ipl.pt

Abstract: This study combines Visible Light Communication (VLC) and Artificial Intelligence (AI) to enhance traffic signal control, reduce congestion, and improve safety, through real-time monitoring and dynamic traffic management. Leveraging VLC technology, the system uses existing road infrastructure to transmit live data on vehicle and pedestrian positions, speeds, and queues. AI agents, employing Deep Reinforcement Learning (DRL), process this data to manage traffic flows dynamically, applying anti-bottleneck and rerouting techniques to balance pedestrian and vehicle waiting times. A centralized global agent coordinates the local agents controlling each intersection, enabling indirect communication and data sharing to train a unified DRL model. This model makes real-time adjustments to traffic light phases, utilizing a queue/request/response system for adaptive intersection management. Tested using simulations and real-world trials involving standard and rerouting scenarios, the approach demonstrates significantly better performance in regard to the rerouting configuration, reducing congestion and enhancing traffic flow and pedestrian safety. Scalable and adaptable to various intersection types, including four-way, T-intersections, and roundabouts, the system's efficacy is validated using the SUMO urban mobility simulator, resulting in notable reductions to travel and waiting times for both vehicles and pedestrians.

Keywords: visible light communication (VLC); deep reinforcement learning (DRL); traffic signal optimization; connected vehicles (CVs); multi-agent systems; urban traffic management; queue/request/response methodology; traffic flow simulation

Citation: Vieira, M.; Galvão, G.; Vieira, M.A.; Véstias, M.; Vieira, P.; Louro, P. Integration of Visible Light Communication, Artificial Intelligence, and Rerouting Strategies for Enhanced Urban Traffic Management. *Vehicles* **2024**, *6*, 2106–2132. https://doi.org/10.3390/vehicles6040103

Academic Editors: Tai-Jin Song and Yao Cheng

Received: 17 November 2024
Revised: 2 December 2024
Accepted: 6 December 2024
Published: 11 December 2024

1. Introduction

Urban traffic management is constrained by increasing vehicle and pedestrian volumes, resulting in congestion, delays, and increased safety risks [1]. With the expansion of road infrastructure no longer a viable option, optimizing traffic flow at intersections has become critical [2]. Adaptive traffic signal control systems, which utilize real-time data, such as data on traffic flows and vehicle queues, provide a promising solution to alleviate congestion and improve the overall efficiency of such systems.

Intersections, key nodes in road networks, frequently become bottlenecks, highlighting the need for intelligent signal control to enhance traffic flow. Advances in technology, especially Deep Reinforcement Learning (DRL), have proven effective in dynamically optimizing traffic signals for vehicles and pedestrians [3,4]. Optimizing traffic flows across multiple intersections is challenging due to the varying conditions involved and the need for information sharing. As intersections are major bottlenecks, intelligent signal control is

vital to reduce congestion. While DRL has shown promise in dynamically managing traffic signals, its application in regard to complex road networks remains difficult. This study addresses these challenges, by integrating DRL with connected vehicle technologies and Visible Light Communication (VLC) to improve traffic management efficiency.

Connected vehicles (CVs) present significant potential in regard to traffic management, by leveraging advanced communication technologies to share real-time traffic and safety information with other vehicles and infrastructure. This connected environment enhances road safety, comfort, and efficiency, enabling the improved optimization of both vehicular and pedestrian flows.

Visible Light Communication (VLC) is another innovative solution that can complement CV technologies to optimize traffic management [5,6]. VLC utilizes the intensity modulation of Light Emitting Diodes (LEDs), widely present in streetlights, traffic signals, and vehicle headlights, to enable seamless data communication within existing infrastructure. This dual-purpose of LEDs, for both illumination and communication, positions VLC as a key enabler in optimizing traffic signals and vehicle trajectories at urban intersections [7,8]. By leveraging LED technology in both road infrastructure and vehicles, VLC paves the way for smarter, more efficient traffic management.

> *"How can Deep Reinforcement Learning (DRL) be effectively applied to optimize traffic signal control and vehicle trajectories at urban intersections, leveraging connected vehicle technologies and Visible Light Communication (VLC) to enhance coordination and reduce traffic congestion?"*

This question addresses the core focus of applying DRL in regard to vehicular communication, specifically to optimize traffic control through the integration of emerging technologies, such as VLC and CVs. By exploring these advanced technologies, this research aims to demonstrate how they can improve traffic flow and intersection efficiency in real-world scenarios.

This paper is organized as follows: After the introduction, Section 2 provides an overview of the complexities and challenges in regard to managing traffic at arterial intersections, focusing on issues such as congestion, signal coordination, and safety. It also discusses the integration of Vehicular Visible Light Communication (V-VLC) for innovative traffic solutions and explains how Reinforcement Learning (RL) can be applied for adaptive, real-time traffic signal control to optimize traffic flows and reduce delays. Section 3 examines the specifics of traffic control in regard to multi-intersection networks, detailing the key components and operational mechanisms. Section 4 introduces a model for traffic signal control, incorporating machine learning techniques and dynamic traffic management. Section 5 presents the simulation results, evaluating the effectiveness of the proposed neural network models in regard to various traffic scenarios. Finally, the conclusion in Section 6 summarizes the key findings, outlines the study's limitations, and suggests potential directions for future research.

2. Background and Literature Review

2.1. Urban Traffic Management and the Challenges

Urban traffic management faces growing challenges from congestion, delays, and safety risks. As cities grow, increasing vehicle volumes strain road networks, particularly during peak hours, with intersections often becoming critical bottlenecks. Factors, such as limited infrastructure, accidents, and outdated traffic signals, worsen the situation, leading to prolonged delays. Moreover, conflicts between pedestrians, cyclists, public transport, and vehicles in shared spaces, further hinder traffic flows in densely populated areas [9].

Traffic delays are heavily linked to congestion, often worsened by static traffic signals that fail to adapt to real-time conditions, causing unnecessary waiting, even on clear roads. During peak hours, road capacity is frequently exceeded, leading to long waiting times and slow travel speeds. A key issue is the lack of communication between adjacent intersections, which hinders traffic signal synchronization and further disrupts vehicle flows.

Traditional traffic management methods, like road expansion and static signals, are limited in regard to their ability to address the complexity of modern urban traffic, as they are reactive, inflexible, and lack sustainability. Emerging technologies [10], such as adaptive traffic control systems, connected vehicles, and intelligent communication systems, offer more effective solutions. However, the current algorithms largely prioritize vehicular traffic, while neglecting pedestrian and cyclist dynamics at intersections. To bridge this gap, reinforcement learning-based traffic control systems must integrate pedestrian considerations. The challenges include managing two-way pedestrian flows on sidewalks, disparities in movement patterns between pedestrians and vehicles, and balancing safety with efficiency [11].

In multi-intersection scenarios, a single-agent approach for traffic light control struggles with scalability, leading to the exploration of collaborative mechanisms. The proposed adaptive traffic control strategy aims to respond to real-time traffic demand in V2X environments, utilizing detailed data for improved control and safety functionalities.

Researchers have explored collaborative mechanisms to address this challenge [12–15], incorporating factors such as queue length in neighboring intersections and modeling relationships between these intersections. These efforts aim to achieve a balance between scalability and efficiency in multi-intersection scenarios, acknowledging the need for a more nuanced approach to optimize traffic control. Our adaptive traffic control strategy aims to respond to real-time traffic demand through current and predicted future traffic flow data modeling. Compared with the traffic flow and occupancy information provided by a fixed coil detector in the traditional traffic environment, the adaptive traffic control system in the V2X environment can collect more detailed data, such as on the vehicle position, speed, queuing length, and stopping time [16]. While Vehicle-to-Vehicle (V2V) links are particularly important for safety functionalities, such as pre-crash sensing and forward collision warnings, Infrastructure-to-Vehicle/Pedestrian (I2V/P) links provide CVs and pedestrians with a variety of useful information.

2.2. Innovative Solutions: V-VLC Integration

Emerging solutions, such as adaptive traffic signal control (ATSC) systems [17], connected vehicles (CVs) [18], and reinforcement learning (RL) algorithms, provide dynamic, real-time traffic management by adapting to fluctuating patterns, reducing delays, and enhancing safety. ATSC uses sensor data to optimize signal timings, while CVs and autonomous vehicles (CAVs) enable advancements, like vehicle speed guidance, to minimize queues and delays. However, implementing these technologies faces challenges, including ensuring seamless integration across different systems and requiring substantial investments in infrastructure, hardware, and software for large-scale deployments.

Advancements in wireless communication and V2V and V2I systems offer opportunities to optimize urban traffic networks by integrating traffic signal control systems with driving behaviors [19]. This paper presents a novel approach that combines VLC localization services with learning-based traffic signal control to achieve comprehensive management of both pedestrian and vehicular traffic. The goal is to reduce waiting times and improve overall traffic safety in multi-intersection scenarios.

To develop an intelligent control system model that facilitates safe vehicle management through intersections using V2V, V/P2I, and I2V/P communications, RL concepts are utilized. RL is a training method that involves rewarding desired behaviors and/or punishing undesired ones [20]. To assess the effectiveness of the proposed V-VLC system in multi-intersection scenarios, we utilize a simulator, the Simulation of Urban MObility (SUMO) [21], which involves simulations that are agent based. As the agent gains experience, it learns to avoid negative situations and focus on positive ones. The traffic lights in the SUMO are controlled by the learning agent, based on its decisions, and the overall flow of traffic is described, while rewarding the actions of the traffic light control agent. The agent's goal is to explore new states, while maximizing the total rewards, to develop the

best possible policy. A dynamic phase diagram and a matrix of states, based on the total accumulated time, are presented to illustrate the concept.

2.3. Connected Vehicles (CVs) and Visible Light Communication (VLC)

CV technology is transforming urban traffic management by enabling communication between vehicles (V2V) and with infrastructure (V2I). This allows for real-time data exchange, such as data on the vehicle speed, location, and road conditions, improving traffic flow, safety, and reducing accidents. CV technology is essential for smart traffic systems that aim to optimize road usage and minimize congestion.

Similarly, VLC is emerging as an innovative solution for traffic management. By modulating the intensity of LED lights, which are commonly used in streetlights and traffic signals, VLC can transmit data, while still serving its primary purpose of illumination. This dual-purpose technology integrates easily into existing urban infrastructure, turning everyday light sources into communication devices.

VLC enhances traffic flow by enabling real-time communication between vehicles and traffic control systems. Its high-speed, localized communication can provide precise, up-to-the-second information to optimize traffic signals, reduce delays, and improve safety. The V-VLC system, featuring a mesh cellular hybrid structure, uses two controllers: a "mesh" controller of streetlights to relay messages to vehicles and a "mesh/cellular" hybrid controller of traffic lights for edge computing [22,23]. In an outdoor environment, various challenges arise. There can be different weather changes, such as rain, fog, snow, and interference from solar or artefactual lights. With the evolution of photodetectors, light filters, signal emitters, and LEDs, many of these obstacles are overcome and there is no major interference in the communication taking place. In extreme cases, where communication via VLC is not feasible, it is guaranteed via RF, as both technologies complement each other.

The system consists of a transmitter generating modulated light and a PIN–PIN based receiver detecting light variations, both of which are wirelessly connected. In Figure 1b, the emitter and receivers' relative position and an illustration of the coverage map, with the footprint regions in the unit cell (#1–#9) and the steering angles δ, are presented as a draft. LED-produced light is modulated using ON–OFF keying (OOK) amplitude modulation. Square unit cells in the environment feature tetra-chromatic white light (WLED) sources at cell corners. The white WLED sources consist of red (R: 626 nm), green (G: 530 nm), blue (B: 470 nm), and violet (V: 390 nm) chips, and combine the lights in the correct proportion to generate white light [24]. Each of the RGBV signals that are sent has a wavelength calibrated amplitude that defines it. Since each VLC infrastructure has four independent emitters, the optical signal generated by the receiver can have one, two (#3, #5, #7, #9), three (#2, #4, #6, #8), or even four (#1) optical excitations, resulting in 2^4 different optical combinations and 16 different photocurrent levels at the photodetector. Filtering is accomplished using a PIN–PIN demultiplexer [22]. The PIN–PIN demultiplexer plays a crucial role in the decoding process, ensuring the accurate retrieval of the original message. It receives the combined OOK signals and, armed with prior knowledge on the calibrated amplitudes, decodes the sent message [25].

The proposed architecture enables Infrastructure-to-Cloud communication (I2IM) through embedded computing platforms for processing and sensor interfacing. It also facilitates peer-to-peer communication (V2V) among vehicles, enhancing data sharing. Streetlights serve as geo-transmitters, strategically placed 20 m apart. This arrangement ensures that the streetlights form a square configuration, which is essential for accurately identifying and illuminating the footprints within the target area. While the general guideline suggests a minimum distance of 30 m, the selected spacing was optimized to align with the geometric requirements and objectives of the study. Each LED transmitter emits I2V messages, including synchronization, physical ID, and traffic information. Upon entering a streetlight's capture range, a unique ID ($q_i(x,y,t)$) is assigned to probe vehicles or pedestrians, along with relevant traffic information.

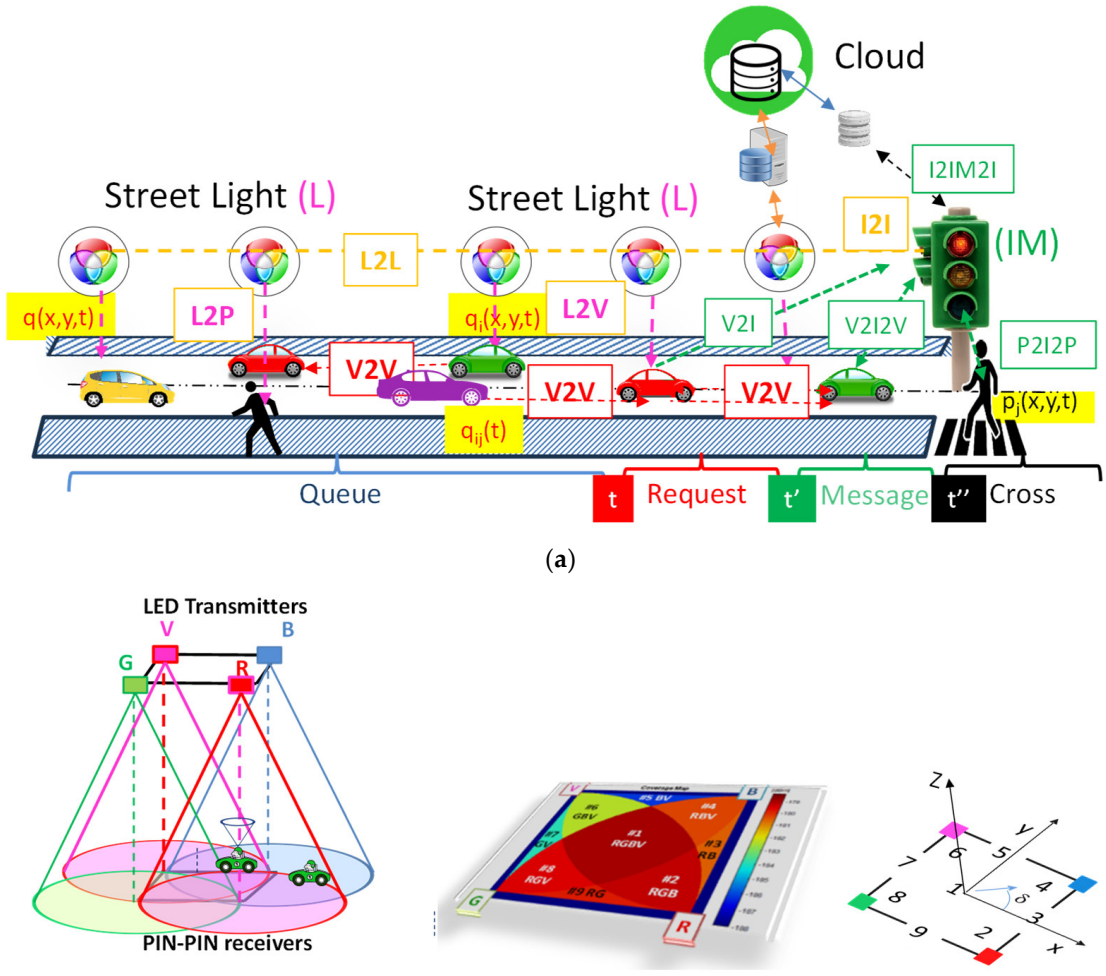


Figure 1. (**a**) A 2D representation of the V-VLC architecture. (**b**) V-VLC emitter and receivers' relative position and an illustration of the coverage map, with the footprint regions in the unit cell (#1–#9) and the steering angle codes (2–9) [22].

The system employs queue/request/response mechanisms and temporal/space relative pose concepts to manage a vehicle's passage through the intersections. Approaching an intersection triggers a crossing request (V/P2I) from the vehicle or pedestrian. The traffic signal responds with an acknowledgment (I2V/P) and the vehicle must follow specified occupancy trajectories denoted as footprint regions (Figure 1b). Caution is exercised if a potential collision risk exists, delaying the response until it is mitigated.

The vehicle speed is calculated using transmitter IDs for tracking as $q_i(x,y,t)$ and mesh nodes estimate indirect V2V relative poses ($q_{ij}(t)$) in scenarios with multiple neighboring vehicles [26]. Requests include positions, directions, and speeds, with leader–follower information aiding subsequent V2I request confirmations. Delays are determined by observing the number of vehicles queuing in each cell at the beginning and end of the green time through V2V2I, as shown in Figure 1a.

The integration of VLC enables direct monitoring among pedestrians, vehicles, and infrastructure, focusing on critical aspects, such as queue formation and pedestrian density, to enhance road safety. P2I2P communication enables travel time calculations, while real-time data on the speed of traffic and waiting times are analyzed using transmitter tracking IDs.

3. Traffic Signal Control at Multi-Intersection Networks

Arterial traffic signal control refers to managing intersections formed by the crossing of two or more main roads, either radial or circular in design. The layout and spacing between intersections vary depending on the 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. Traffic Scenario and Phase Diagram

The traffic scenario analyzed in this study is depicted in Figure 2a. It includes three uniform 4-arm junctions, spaced at varying distances. The lane between junctions C0 and C1 is 400 m long, while the lane between C1 and C2 is only 200 m long. Figure 2b presents the simulated environment for each four-legged intersection, showcasing the optical infrastructure (Xij), the generated footprints (#1–#9), and the interactions between connected vehicles and pedestrians. The streetlights along the lanes, denoted as Xi,j, are identified by integers that follow the opposite direction of traffic (N, S, E, W), starting from *streetlight 0* at the required signalized intersection and extending, in line, towards *streetlights K*, *L*, *M*, and *N* at the neighboring junction.

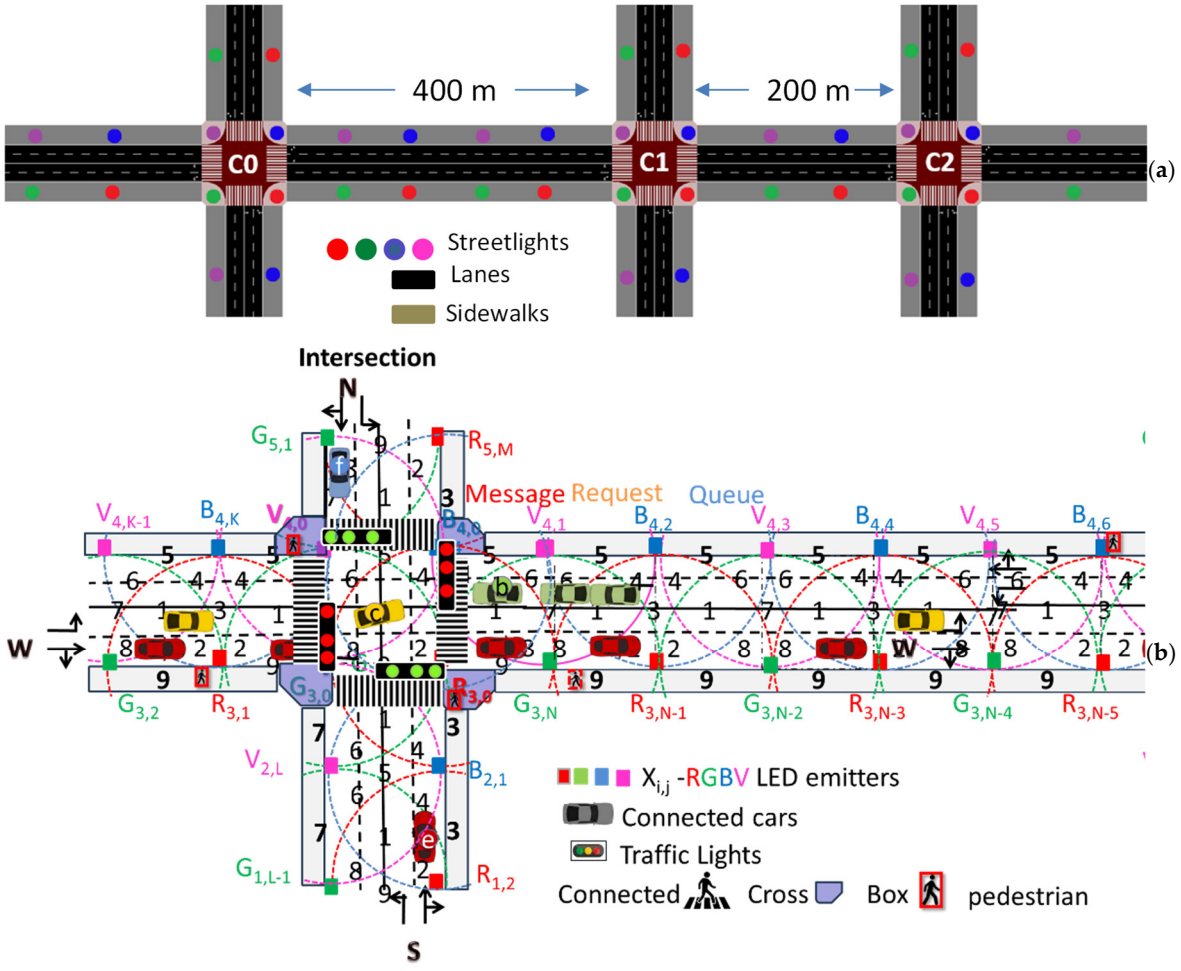
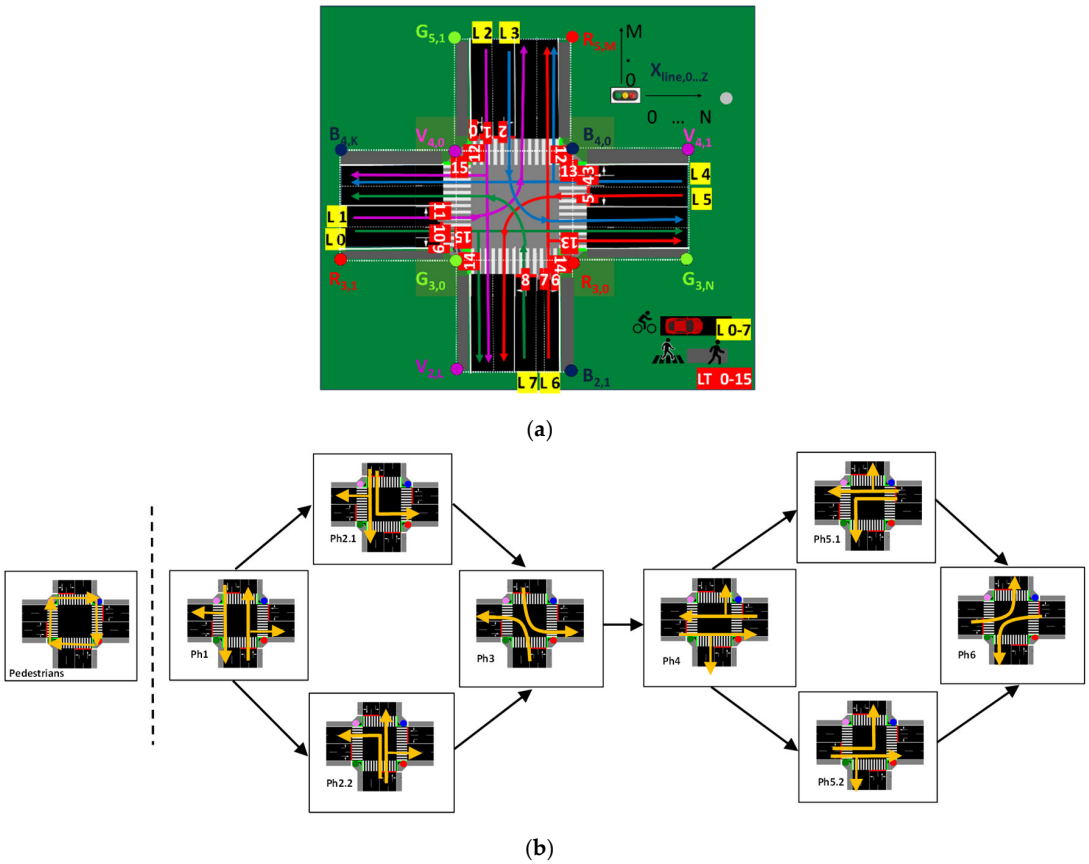


Figure 2. (**a**) Environment scenario; (**b**) simulated scenario for each junction: four-legged intersection and an environment with the optical infrastructure (X_{ij}), the generated footprints (#1–#9), and the connected cars and pedestrians.

Each arm of the junction features two lanes: one designated solely for left turns and the other allowing vehicles to either continue straight or turn right. Using VLC, as shown in Figure 1, the vehicle gathers information about the footprint regions along the junctions, their respective lanes, and the surroundings. This data helps the vehicle determine the appropriate lane for its intended route. Once correctly positioned, the vehicle communicates with the Intersection Manager (IM) via V2I technology to signal its intention to proceed in the specified direction. Additionally, the environment includes sidewalks for pedestrian movement, with designated waiting zones at each intersection. These zones serve as safe waiting areas for pedestrians until their crossing phase is activated, allowing them to safely use the zebra crossings.

For vehicles, there is a north–south phase, where they can either proceed straight forwards or turn right, followed by a phase where vehicles coming from the north can cross the intersection in all directions, and another phase where vehicles from the south can do the same. Additionally, there is a phase where vehicles traveling from both the north and south can only turn left. The same phase structure applies to the east–west direction. Pedestrians have an exclusive phase during which all vehicle traffic lights are red, allowing them to cross the intersection safely without any interference from vehicles. This exclusive phase ensures pedestrian safety by preventing any crossover between pedestrians and vehicles at crossings.

The signal timing is determined by factors such as traffic demand, intersection layout, and traffic management objectives (e.g., minimizing delays, maximizing traffic flow). Traffic signals typically operate in phases, with green lights allowing movement in one direction, while red lights restrict conflicting movements. In this study, we designed eight vehicle signal phases and one exclusive pedestrian phase for each intersection, as illustrated in Figure 3b.



(**a**)

(**b**)

Figure 3. (**a**) Schematic diagram of one junction with coded lanes (L/0–7) and traffic lights (TL/0–15). (**b**) Phase diagram with the traffic directions [22].

A traffic control system consisting of sixteen traffic lights (TLs) has been implemented to manage the flow of vehicles approaching the intersections. These traffic lights are numbered (TL 0–15), as shown in Figure 3a, which also displays the numbering of the lanes (L 0–7), which is consistent across all three junctions. These sixteen traffic lights enable the implementation of traffic phases to regulate the flow of vehicles and pedestrians, as illustrated in Figure 3b.

The main challenge in controlling traffic across multiple intersections is the coordination required to manage the varying traffic conditions between intersections. The target roads between intersections vary in regard to their distance, making it crucial to synchronize traffic signals based on the traffic volume, vehicle movements, and road capacity. Without proper coordination, localized traffic signal control can lead to inefficiencies, like congestion or increased waiting times, especially when multiple intersections interact closely in an arterial network.

3.2. Integration of VLC and CV Technologies

The communication protocol defines the structure and rules governing the exchange of information. It includes specifications for the synchronization, identification, and payload portions of the transmitted frame. The communication protocol is presented in Table 1.

Table 1. Message protocol defined for each of the V-VLC communications.

	SOF 5 bits	**TIME** 6+6+6 bits			**Flag** 3 bits	**COM** 4 bits	**POSITION** 4+4 bits		**PAYLOAD** 4+4+4+4+4+4 bits					
L2D	Sync	Hour	Min	Sec	END	1	y	x	0000+0000					EOF
V2V	Sync	Hour	Min	Sec	END	2	y	x	Lane (0–7)	Device (nr)	Device IDy	Device IDx	Nr. behind	…. EOF
V2I	Sync	Hour	Min	Sec	END	3	y	x	TL (0–15)	Device (nr).	Device IDy	Device IDx	Nr. behind	…... EOF
I2V	Sync	Hour	Min	Sec	END	4	y	x	TL (0–15)	Device ID	Device IDy	Device IDx	Nr. behind	…. EOF
P2I	Sync	Hour	Min	Sec	END	5	y	x	TL (0–15)	N, S, E, W.	…..…			EOF
I2P	Sync	Hour	Min	Sec	END	6	y	x	TL (0–15)	Phase	…..…			EOF

The communication protocol starts with a frame structure that begins with a 5-bit synchronization block (Sync) represented by the pattern [10101]. This block is used to synchronize the receivers and marks the start of a new frame (SOF). Next, the timeline information is provided in the TIME block, consisting of a 12-bit sequence (6 + 6 + 6) that encodes the hour, minute, and second. After this, a flag with the pattern [1111] alerts the decoder to expect specific ID blocks. All are 4-bit blocks beginning with the communication type (COM), which specifies the type of communication between the streetlights (L), vehicles (V), pedestrians (P), and infrastructure (I). Next comes the localization of the transmitters, defined by y and x coordinates and, depending on the communication type, the frame also includes details about the occupied lane (Lane 0–7), the traffic light signals requested (TL 0–15), the number of vehicles behind the leader (Veic. nr), the ID assigned by the Intersection Manager (ID) to acknowledge vehicle messages, the cardinal direction (Direct.), and the active phase (Phase), all provided by the infrastructure through either a "request" or "response" message at the intersection.

For traffic-related messages, these blocks carry vehicle information, such as y and x coordinates, the position of the cars behind the leader (CarIDy, CarIDx, and the number of following vehicles), as well as traffic-related data (Payload), like the road conditions, average waiting times, and weather information. The frame concludes with a 4-bit End of Frame (EoF) block, represented by the pattern [0000], which signals the end of the frame.

In Figure 4, the first moments to be studied for both the vehicle and the pedestrian are pointed out. The highlighted car is coming from the north, heading towards C0, and is waiting for its phase to be active, so it can turn left. At this moment, various VLCs will be

studied, including V2V, V2I, and I2V. In C0, the pedestrians in the designated waiting area are waiting for the pedestrian phase to be activated. During this time, P2I communications are established, where a request to cross is made, and I2P serves as the IM's response.

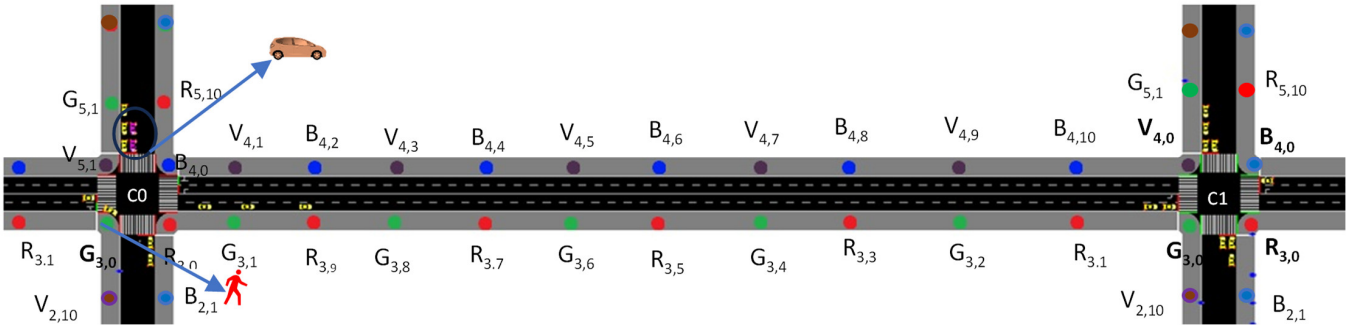


Figure 4. Simulated VLC in a two junction (C0 and C1) scenario, involving RGBV ID transmitters.

Figure 5a and Figure 5b demonstrate the MUX signal and the decoded messages between the vehicles or the pedestrians and the traffic lights, respectively.

In Figure 5a, for the V2V communication (COM: 2), the vehicle behind the target car communicates with the one in front, providing its position as $G_{5,1}$, $R_{5,10}$, $V_{4,0}$, the lane it is in (Lane: 2), the number of vehicles following it (in this case, none), the time of communication (11:18:30), and, as there are no cars behind it, it does not transmit anything in these blocks to the car in front. After receiving this communication, the leader then makes a request to the IM through V2I. It provides its position as $G_{5,1}$, $R_{5,10}$, $V_{4,0}$, the traffic light to which the request is being made (TL: 2), the number of vehicles following it (Veic. (nr): 1), the time of the communication (TIME: 11:18:31), the car identifier (y, x: $G_{5,1}$), and the number of cars behind the follower, which is currently 0. Next, the I2V communication occurs, where the leader receives a response with the same information at 11:18:32, indicating that the active phase is 4 (W > E). After the N > S Left phase is activated, the cars move toward the C1 intersection, where they are currently lined up waiting.

In Figure 5b, for the P2I communication, the pedestrian's position (y, x: $G_{3,0}$ $V_{2,10}$) is transmitted, along with the traffic light (TL: 14), the intended direction (in this case, east, indicated as 3), and the communication time of 15:35:20. For the I2P response from the IM, the position, $G_{3,0}$ $V_{2,10}$, and traffic light (TL: 14), the currently active phase (in this case, NS Left, phase 3), and the communication time of 15:35:21 are sent. After crossing intersection C0, the pedestrian eventually reaches another waiting area, this time at intersection C1, where P2I and I2P communications are established again.

Here, for P2I, the pedestrian's position $G_{3,0}$ $V_{2,10}$ is transmitted, along with TL 14, the intended direction (east, 3), and the communication time of 15:40:47, which are also assigned. For the I2P response from the IM, its ID ($G_{3,0}$ $V_{2,10}$, TL 14), the currently active phase (WE, phase 4), and the communication time of 15:40:48 are sent. After crossing intersection C1, the pedestrian reaches C2, where, after some time, they may wish to cross again, in the same or in another direction (P_{C2}). Finally, for 2V, the last P2I C2 and I2P C2 messages are exchanged.

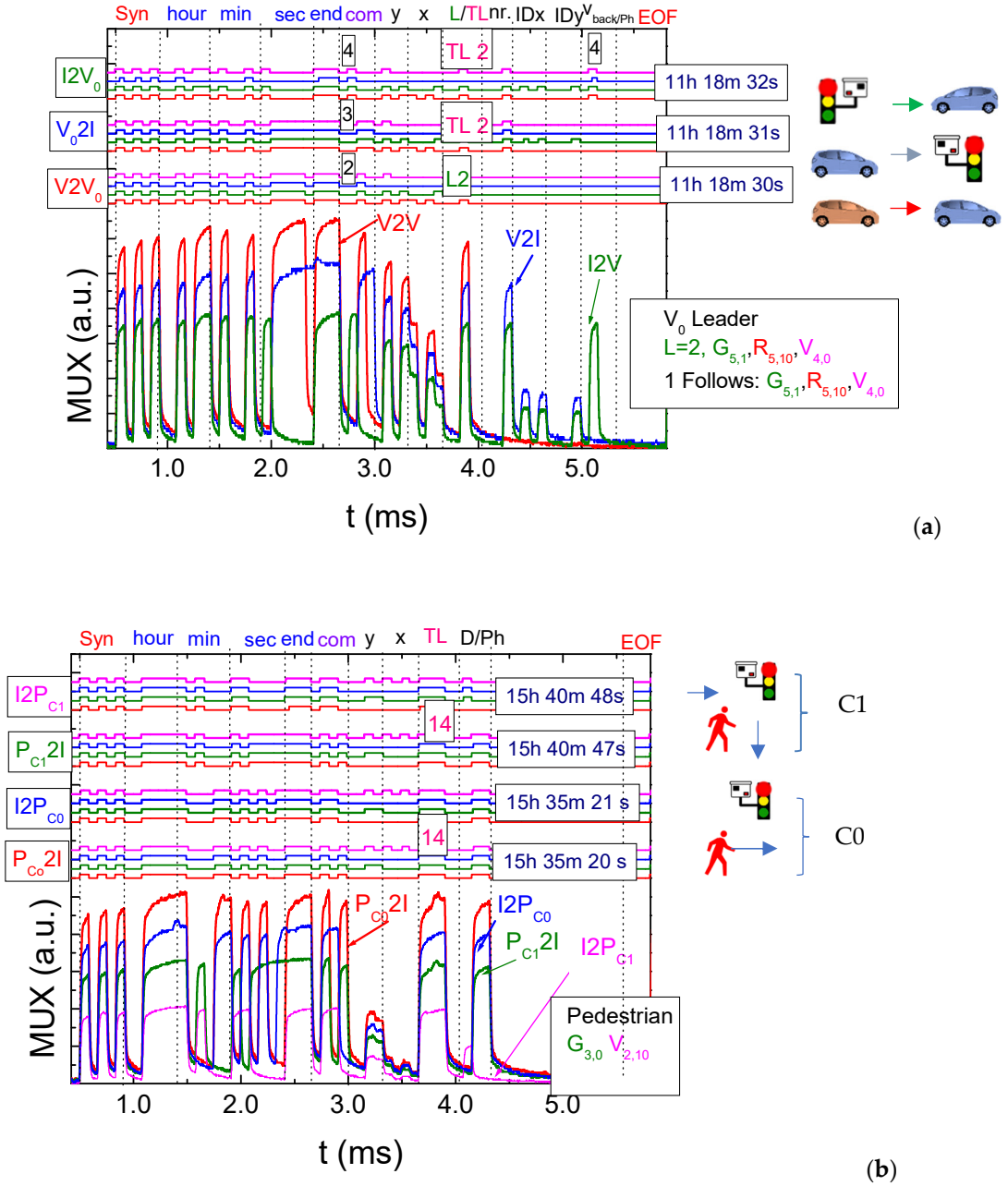


Figure 5. Normalized MUX signal responses and the corresponding decoded messages, displayed at the top, sent by the IM to: (**a**) the vehicles. (**b**) pedestrians waiting at the corners (I2P1,2) for various frame times. On the right-hand side, the analyzed communication type is displayed to assist visual interpretation.

The results show that it is possible to use VLC to detail the flow of Vehicle-to-Infrastructure, Vehicle-to-Vehicle, Pedestrian-to-Infrastructure, and Infrastructure-to-Pedestrian communications at various intersections, illustrating a structured communication framework for coordinating traffic and pedestrian movement, as follows:

1. *V2V and V2I communication:* Vehicles communicate both with one another and with the infrastructure to relay information on vehicle positioning, traffic light phases, and vehicle status. This data exchange helps vehicles align their movements with active traffic phases, ensuring a coordinated flow;
2. *P2I and I2P communication*: Pedestrians participate in a similar communication cycle, requesting to cross intersections and receiving confirmations. The infrastructure

responds based on the active traffic phase, managing pedestrian crossings in harmony with vehicular phases, to improve safety and efficiency;

3. *Traffic flow and phase coordination:* Each intersection has specific phases (Figure 3) to control pedestrian and vehicle movements. By synchronizing these phases across multiple intersections (C0, C1, C2), the system can handle complex flows and enhance safety, preventing conflicts between vehicles and pedestrians.

Overall, the V-VLC protocols provide a structured way for traffic management systems to coordinate and streamline both pedestrian and vehicle movements, which can be particularly beneficial in densely populated or high-traffic areas. The approach leverages real-time data to reduce delays, improve safety, and optimize intersection efficiency.

3.3. DRL Framework for Traffic Signal Control

To manage the data collected via VLC, a system is needed to intelligently control traffic and respond to vehicle maneuvering requests in real-time [27–29]. For this, a Multi-agent Reinforcement Learning (MARL) system is implemented. A scenario with three connected four-arm intersections, each with two lanes in both directions, is considered (Figure 2a). Each intersection is controlled by an agent that maps the surrounding environment, acquiring information about vehicles and pedestrians. Communication between the infrastructure and vehicles enables a cooperative information-sharing environment. The traffic data collected by the three agents is stored and used to train a single network that controls the entire environment by selecting the best phase for each intersection.

A Deep-Q network (DQN) is used, trained using the deep Q-Learning technique. The Q-value represents the expected cumulative reward of taking a certain action in a state and following the optimal policy thereafter. In traditional Q-Learning, the Q-value for each state–action pair is stored in a look-up table, known as tabular Q-Learning. This method guarantees convergence between infinite visits and state–action pairs, but it is only effective for small-scale problems. For large-scale, continuous state–action spaces, deep Q-Learning networks are used, where a neural network predicts the Q-values based on the input state. Figure 6 illustrates how the intelligent system works during a simulation and training.

At each time step t, the network receives a state input st, based on the observation of the environment by the agent and then executes an action at, which transforms the state observed to the next state, st+1. Then, the reward, rt, a metric that defines how good the action was for the environment, is calculated. The reward used considers both vehicle and pedestrian waiting times, as can be seen in Equation (1). Moreover, $wt_{veh,t}/wt_{ped,t}$ is the amount of time in seconds a vehicle or a pedestrian has a speed of less than 0.1 m/s at t, since they arrive in the environment. Additionally, n represents the total number of vehicles/pedestrians in the environment at t. With this metric, the values of $atwt_t$ do not reset, until the vehicle or pedestrian crosses the intersection.

$$atwt_{veh,t} = \sum_{veh=1}^{n} wt_{(veh,t)} \qquad atwt_{ped,\,t} = \sum_{ped=1}^{n} wt_{(ped,t)} \tag{1}$$

The final reward equation, r_t, is defined in Equation (2); where $atwt_t$ and $atwt_{t-1}$ are the accumulated total waiting time of all the cars/pedestrians in the intersection captured, respectively, at t and $t-1$. The weights of p_{veh} and p_{ped} are set based on the desired priority that the agent should have in terms of the vehicles and pedestrians during network training. The agent will learn a policy that benefits one more than the other, or will keep the system balanced, if the weights are equal.

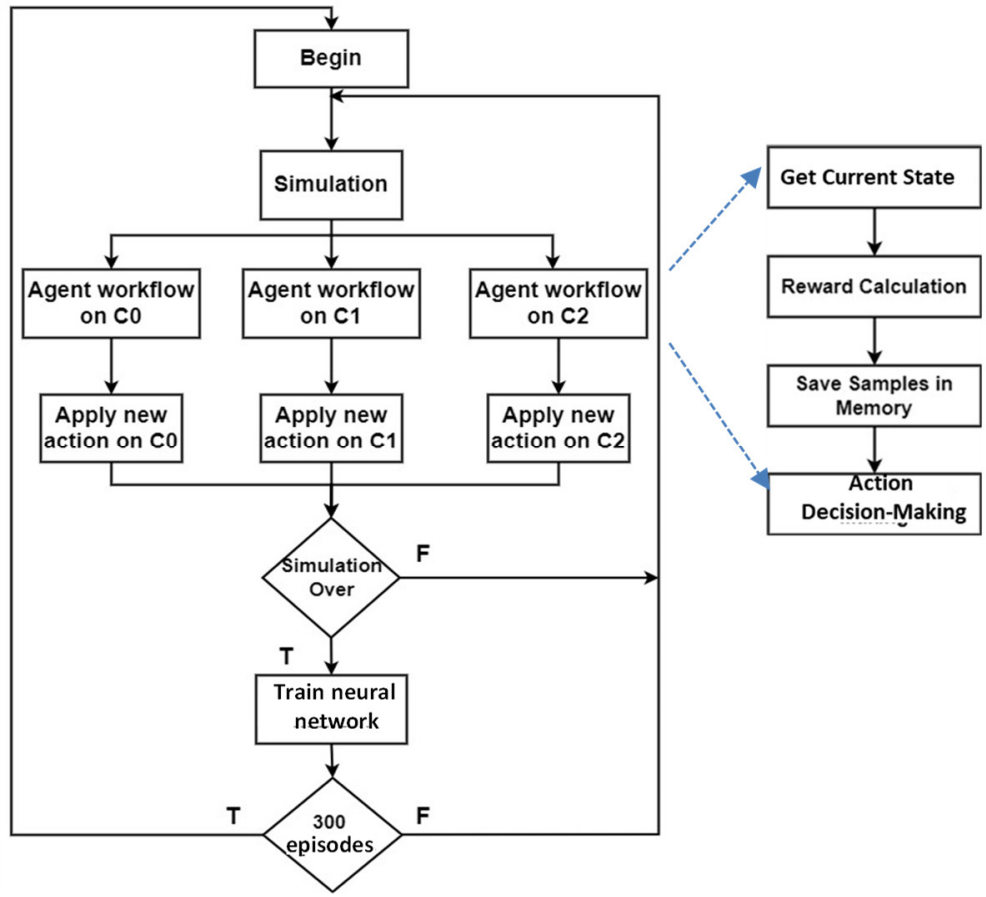


Figure 6. Flowchart during simulation and training.

$$r_t = p_{veh}(\text{atwt}_{veh,t-1} - \text{atwt}_{veh,t}) + p_{ped}\left(\text{atwt}_{ped,\ t-1} - \text{atwt}_{ped,t}\right) \quad (2)$$

This experience ex = (*st*, *at*,*rt*,*st*+1) will be stored in the replay memory, to be used in the future to train the agent. The replay memory is a dataset of an agent's experiences Dt = (e1, e2, . . ., et), which are gathered when the agent interacts with the environment as time goes by (t = 1, 2. . ., n). During training, a mini batch of random samples are chosen. This random selection of samples breaks the temporal correlation between consecutive samples. If the agent learned only from consecutive samples of experiences, as they occurred sequentially in the environment, the samples would be highly correlated and would, therefore, lead to inefficient learning. The replay memory buffer is filled up to a specific frame length and when it is full, old experiences are overwritten by new ones.

The neural network architecture implemented consists of a fully connected layer network (FCLN) and the weights θk in the FCLN are used to approximate the Q-values, Q(s, a; θk). The first layer of the network is the input layer, formed by an input layer of 164 neurons, representing the state of the environment. Following this, there are five hidden layers, each one with 400 neurons, each with a rectified linear unit (ReLU), an activation function commonly used in deep neural networks, with the ability to introduce non-linearity into the network, allowing the NN to learn complex patterns and representations from the data. Finally comes the output layer, with nine neurons, according to which each one will display the Q-values for each action. The next action that the agent will choose is determined by the maximum Q-value output.

To improve this prediction in terms of the Q-values, a Mean Squared Error function is used, which is a mathematical function that quantifies the difference between the predicted Q-values and the target Q-values, as is seen in Equation (1).

$$MSE\ Loss = \frac{1}{N}\sum_{i=1}^{N}\left(Q_{target} - Q_{pred}\right)^2 \quad (3)$$

where N is the number of samples in the memory, Q_{pred} is the Q-value predicted by the main network, and Q_{target} is acquired using a network similar to the main one, but which is not trained. At certain intervals between epochs, episodes that are used to train the main network, the weights of this network are copied to the secondary network, bringing greater stability to the training process. The Q_{target} values are calculated based on Equation (2). Where *rt* is the reward obtained and γ is a discount factor applied to the $maxQ_{target}$ value, lowering the importance of the future reward compared to the immediate reward.

$$Q_{target} = r_t + y.maxQ_{target}(s_{t+1,a}) \quad (4)$$

The MSE loss function calculates the squared difference between the predicted and target values. The goal during training is to minimize this loss, improving the predictions and making the agent's decisions more accurate. The network adjusts the weights of the neurons to approximate the Q-value closer to the target Q-value, with the loss decreasing as the model improves its predictions.

The network is evaluated based on the decisions made, such as the average number of cars in queues. The results show that it can manage traffic at intersections, but without information exchange between agents about lane occupancy, vehicle queues can increase. As traffic environments grow more complex, communication between agents becomes crucial. In simpler scenarios, where vehicles only need to wait at a single intersection, phase control is straightforward. However, with multiple junctions, vehicles may not leave the environment, but rather move to another intersection, making coordination and information exchange between agents more important. In cases where intersections are close or traffic increases during rush hour, effective communication between agents can help avoid long queues and waiting times by adjusting phase activation strategies based on both local and neighboring conditions.

4. Proposed Approach and Methodology

Different DRL methods can be used, depending on the traffic scenario. For multiple intersections, these methods are classified into centralized and decentralized control approaches.

4.1. Multi-Agent Reinforcement Learning (MARL) Systems

In this case, the implemented method is centralized, as shown in Figure 7, which depicts the algorithm used. Although there is no direct communication between agents, it is assumed that they communicate indirectly by sharing their experiences. Our work has explored different types of junctions, starting with one [23], then evolving to two [21], and now three, where we have closely examined the behavior and dynamics of various traffic scenarios. Since the junctions are homogeneous, the similar experiences observed by each agent at every junction can be used to train a shared neural network. This network then acts as a global agent that makes decisions and determines the actions to take. So far, this approach has proven effective in managing both pedestrian and vehicle traffic. However, there are occasional peaks in queue occupancy at the junctions. High numbers of cars waiting can put pressure on junctions, disrupting smooth traffic flow. The global agent makes decisions based on individual junctions, without considering the queue states at neighboring junctions.

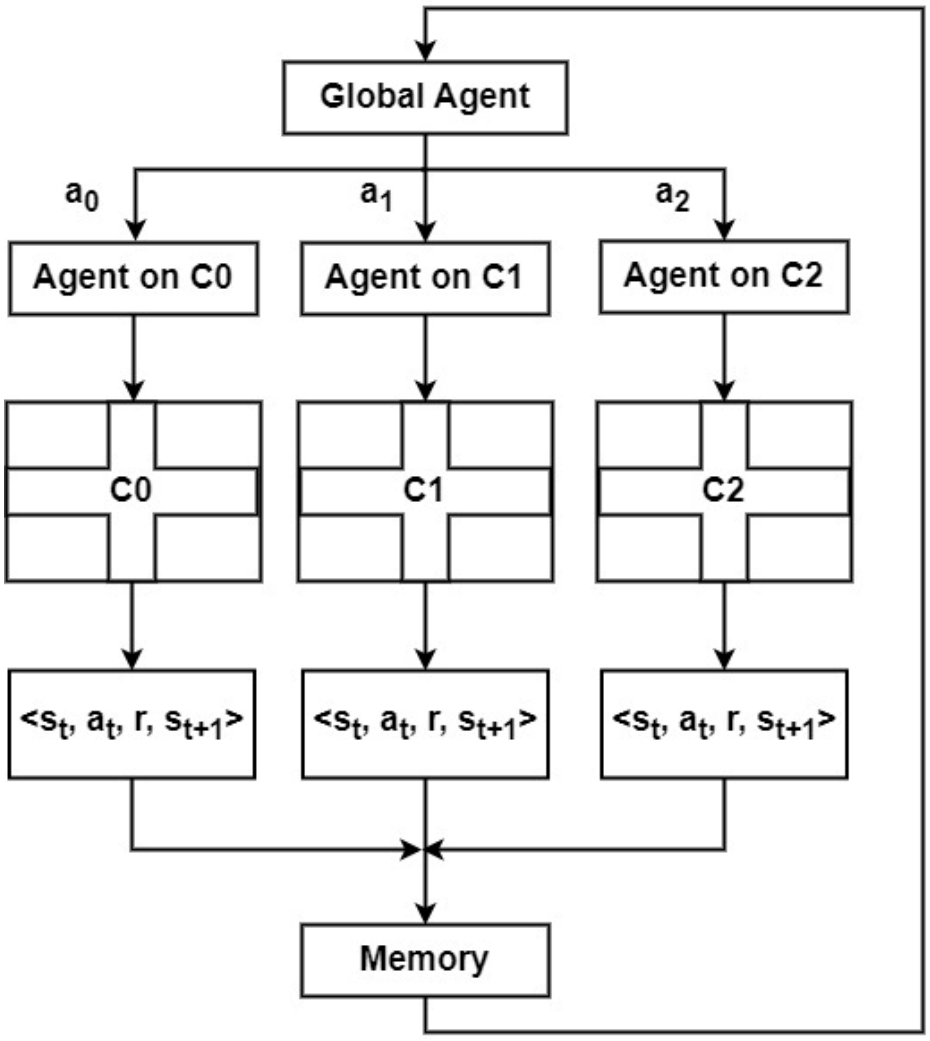


Figure 7. A schematic diagram of the algorithm employed, using centralized MARL.

Based on the study across various scenarios, we gained insights into traffic queuing behavior and identified the capacity limits of each lane. To address the lack of direct communication between agents, we incorporated this prior knowledge into the network by establishing threshold values for the queues. This enables the global system to regulate traffic flows in critical sections, evaluating the volume of traffic that can be accommodated in each direction.

In our work, we started with a typical arterial scenario, referred to as the "standard scenario", where most vehicles (75%) travel straight forward and the remaining 25% turn right or left. However, when traffic demand exceeds the system's capacity, due to incidents or heavy congestion, the system switches to an alternative strategy, the "rerouting scenario". In this scenario, 75% of the vehicles are redirected to turn right or left, while only 25% proceed straight ahead. This strategy helps balance the traffic load when certain sections or chains of sections become congested.

Although the neural network is trained centrally, traffic signal agents at each intersection locally implement the rerouting strategies. The system establishes congestion thresholds, which are used to adapt and optimize traffic flows at the intersections. This ensures efficient traffic management, even in high-demand conditions.

4.2. Standard and Rerouting Traffic Scenarios

For the simulation environment, a three-connected four-arm intersection scenario, with two lanes in both directions, is considered (Figure 2). Each one of these intersections

is controlled by an agent that maps the environment around it in cells, acquiring different information about the vehicles and pedestrians travelling through the intersection (Figure 5). Each of the three intersections is divided into three layers that are made up of 164 cells. The first layer is made up of 80 cells, 10 for each lane routing vehicles to the junction, indicating their presence. If a vehicle is inside the cell, it is filled with '1', otherwise it is filled with '0'. The second layer, made up of the same number of cells, indicates the normalized speed of the cars in each cell, if any are present. The third layer, made up of just four cells, represents the waiting zones, indicating the number of pedestrians standing still, waiting for their phase to become active. This representation state helps the agent to map the environment around the intersection and ends up being very similar to the states observed via VLC, as illustrated in Figure 8. In this case, the vehicles are identified over time by the lane they are in and by the traffic light they are communicating with. Pedestrians, on the other hand, are identified over time by the waiting zone they are in, as well as by the traffic light they are communicating with.

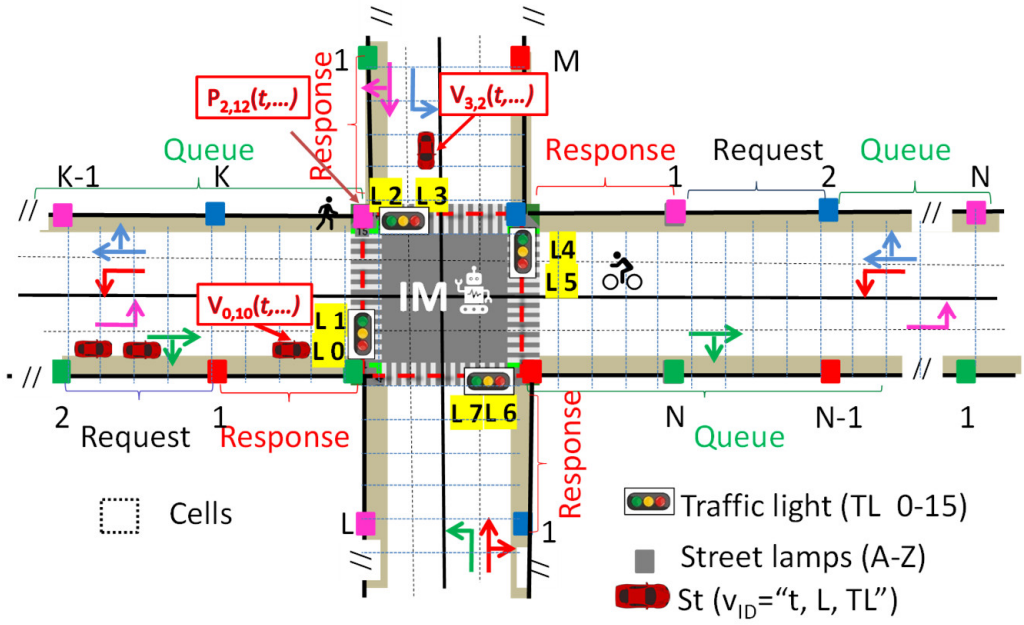


Figure 8. A schematic diagram of the representation state for each junction.

Two traffic scenarios are considered. A scenario where 75% of the vehicles go straight ahead and the remaining 25% turn right or left, called the *standard scenario*. The second scenario, with 75% of vehicles turning right or left and 25% going straight ahead, is called the *rerouting scenario*.

These two scenarios address different traffic conditions: in the arterial scenario, most vehicles move straight through the artery, while in the rerouting scenario, the focus shifts to redistributing the traffic through turning. To improve traffic management in the rerouting scenario, upstream anti-bottleneck and smart rerouting techniques are applied. These methods adjust the intersection control in real-time based on current congestion levels, dynamically assigning priority to alternative routes, using radial or circular driving strategies. The aim is to observe the differences in the traffic environment caused by previously acquired information on queues through studies involving extreme traffic scenarios that are introduced into the network, defining threshold values for them.

5. Results and Discussion

5.1. Network Training

To compare both scenarios, two neural networks were trained with a reward function of 50/50, this being the weight for vehicles and pedestrians, respectively. The environment simulated 2600 vehicles and 2000 pedestrians over 300 episodes, each lasting 3600 s. The traffic environment was simulated in SUMO and the neural networks implemented with the TensorFlow library. All the code involved was implemented in Python, in order to make the connection between the neural networks and the environment. The training parameters are detailed in Table 2.

Table 2. Training parameters.

Parameter	Value
Number of episodes	300
Maximum steps	3600
Vehicles generated	2600
Pedestrians generated	2000
Hidden layers	400
Activation function	ReLU
Width layers	400
Batch size	100
Learning rate	0.001
Training epochs	800
Memory size	50,000
Number of states	164
Number of actions	9
Gamma	0.75

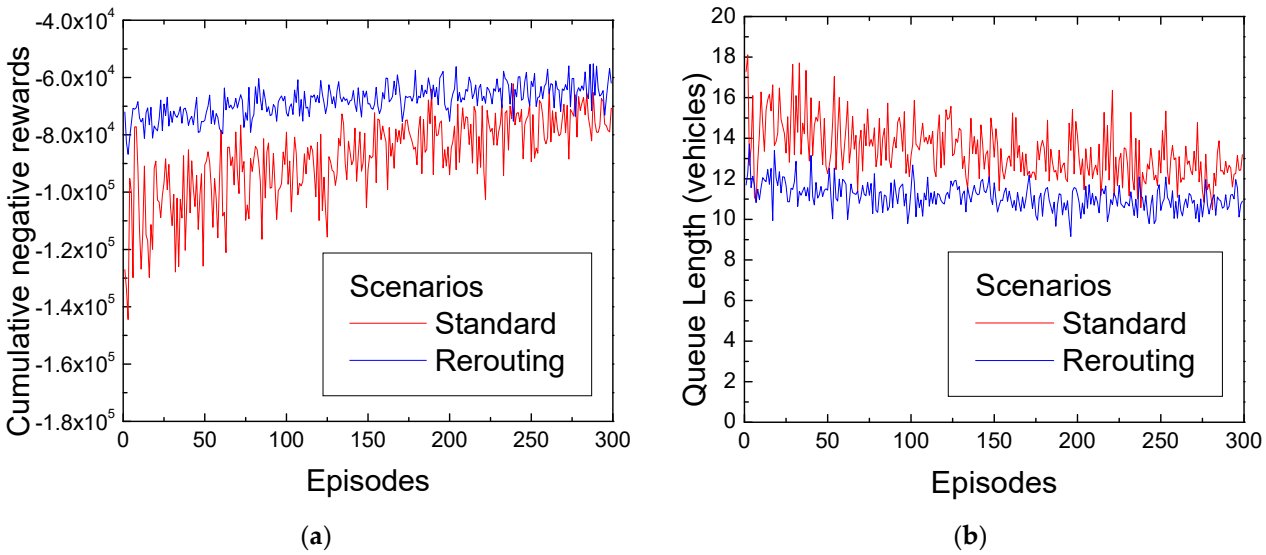


Figure 9. Network training for both scenarios: (**a**) cumulative negative rewards and (**b**) average queue size.

In order to analyze the results obtained, Figure 9a presents a cumulative negative reward graph. A significant difference can be observed between the curves. As training progresses, the cumulative negative reward curves gradually converge, indicating the ability to learn and optimize the control strategies across the standard and rerouting scenarios, stably. In terms of the rerouting scenario, the reward oscillations are smaller compared to the standard scenario. Since a reward is determined by a 50% weighting in terms of vehicle waiting times and a 50% weighting in terms of pedestrian waiting times, the standard scenario consistently experienced greater pressure due to longer queues and waiting times. This accounts for the oscillations in the graph, which were expected.

Figure 9b illustrates the average queue size during training for both scenarios.

Overall, both networks were well-trained, as the rewards became progressively less negative over the episodes. The rerouting reward distribution becomes more consistent than in the standard scenario, suggesting faster learning of optimal strategies and excellent stability, exhibiting robust convergence and effectiveness.

The standard scenario clearly exhibits greater oscillation compared to the rerouting scenario, as expected. This underscores the importance of queue limit awareness in the rerouting scenario. By utilizing predefined threshold values, the system strategically activates certain phases, prioritizing those that reduce pressure on neighboring junctions and promote smoother traffic flows.

5.2. Network Testing

To test the trained networks, simulations were carried out for 3600 s with 2600 vehicles and 2000 pedestrians. The evaluation metrics and comparison validate the optimized performance of the proposed algorithm in terms of traffic safety and efficiency. Additionally, standard and rerouting halting vehicles, and halting pedestrians, are considered to further evaluate the optimization effects.

Figure 10a–c shows the halting vehicles at junctions C0, C1, and C2, respectively. In the scenario with three junctions arranged horizontally, it was determined that C1 is the critical junction. Positioned between C0 and C2, C1 receives traffic from both directions, which puts it under significant pressure, increasing vehicle queues.

By incorporating threshold values related to queue lengths, the system becomes more responsive to the level of congestion on critical roads, particularly those connecting junctions C0/C1 and C1/C2. This results in rerouting vehicles that would otherwise contribute to the congestion. For instance, vehicles coming from the west and heading east at junction C0 are informed of the road conditions between C0 and C1 when they enter the environment. If the number of vehicles on this road exceeds a set threshold, such as 25 cars (even though the 400 m road can hold more), vehicles at C0 will alter their route and turn right instead. This rerouting reduces the queue at junction C0, as shown in Figure 10a. Vehicles that change their route exit the environment by turning right, preventing them from adding pressure on the next junction. This leaves space for new cars to enter, while also avoiding situations where vehicles are stalled at a green light due to a lack of room to move into the next congested lane.

Additionally, the system can manage congestion by activating phases that direct fewer cars into critical lanes. For example, instead of activating the west–east straight phase, which would funnel many cars into C1's critical lane, the system could activate the north–south straight phase, or even a pedestrian phase, if needed. This would allow the cars in C1 to be cleared from the critical lane, creating space for incoming traffic. In Figure 10b, a significant difference in the number of vehicles stopped at the C1 intersection can be observed. Between approximately 800 and 2200 s, the standard scenario shows that the intersection reaches high occupancy levels, indicating heavy congestion. In contrast, during the same time interval, far fewer cars are stopped at C1 in the other scenario.

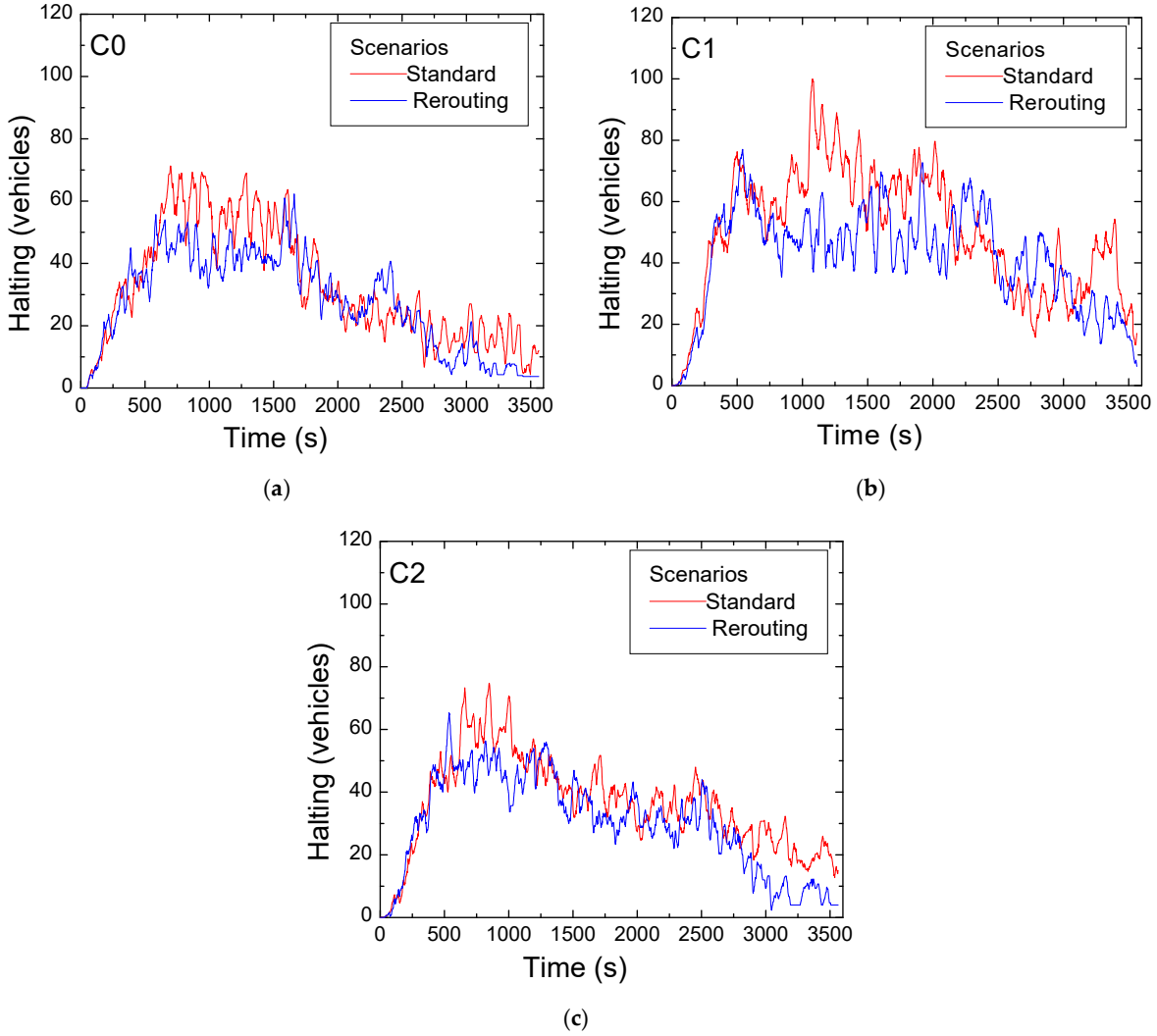


Figure 10. Comparison of trends over time for vehicle halting sessions at intersections in standard versus rerouting scenario: (**a**) intersection C0, (**b**) intersection C1, and (**c**) intersection C2.

As the critical intersection, C1 is situated between C0 and C2, the two main intersections, and is connected to the two most important roads in the network. This makes it crucial for C1 to exercise micro-control over the traffic passing through it, particularly in relation to the length of the road between C0 and C1. With a length of 400 m, this road has a high vehicle flow capacity. Without strict control, this could overwhelm the C2 intersection, as the road connecting C1 and C2 is only 200 m long, half the capacity of the C0–C1 road, and has a defined limit of 10 cars. To prevent this, the system must implement route changes or activate phases that reduce the number of vehicles on these roads.

On the other hand, vehicles traveling from C2 to C1 and heading toward C0 will find it easier to move through the intersection, as they are transitioning from a shorter 200 m road to the 400 m road, which has greater capacity. However, the system must still ensure that the road's limit of 25 cars is respected to maintain a smooth traffic flow.

In Figure 10c, the number of vehicles stopped at the C2 intersection is shown. Over time, the rerouting scenario consistently presents fewer stopped vehicles compared to the standard scenario, as anticipated and discussed earlier. This behavior is similar to what is observed at the C0 intersection, though the management of C2 is slightly more critical due to the route changes and phase activations involved.

Compared to the high queue levels at C0, C1, and C2, the rerouting scenario reduces the average traffic pressure by 66%, 50%, and 75%, respectively. This demonstrates that rerouting significantly improves the efficiency of arterial roads and offers practical benefits.

The road connecting C2 to C1 is only 200 m long, which gives it half the vehicle flow capacity of the other roads in the network. As a result, traffic at C2 must be carefully managed to prevent C1 from becoming overly congested.

In Figure 11, the number of halted pedestrians at each intersection is displayed for both the standard and rerouting scenarios. Comparing pedestrian traffic between these scenarios, the rerouting scenario shows fewer pedestrians waiting in the designated areas.

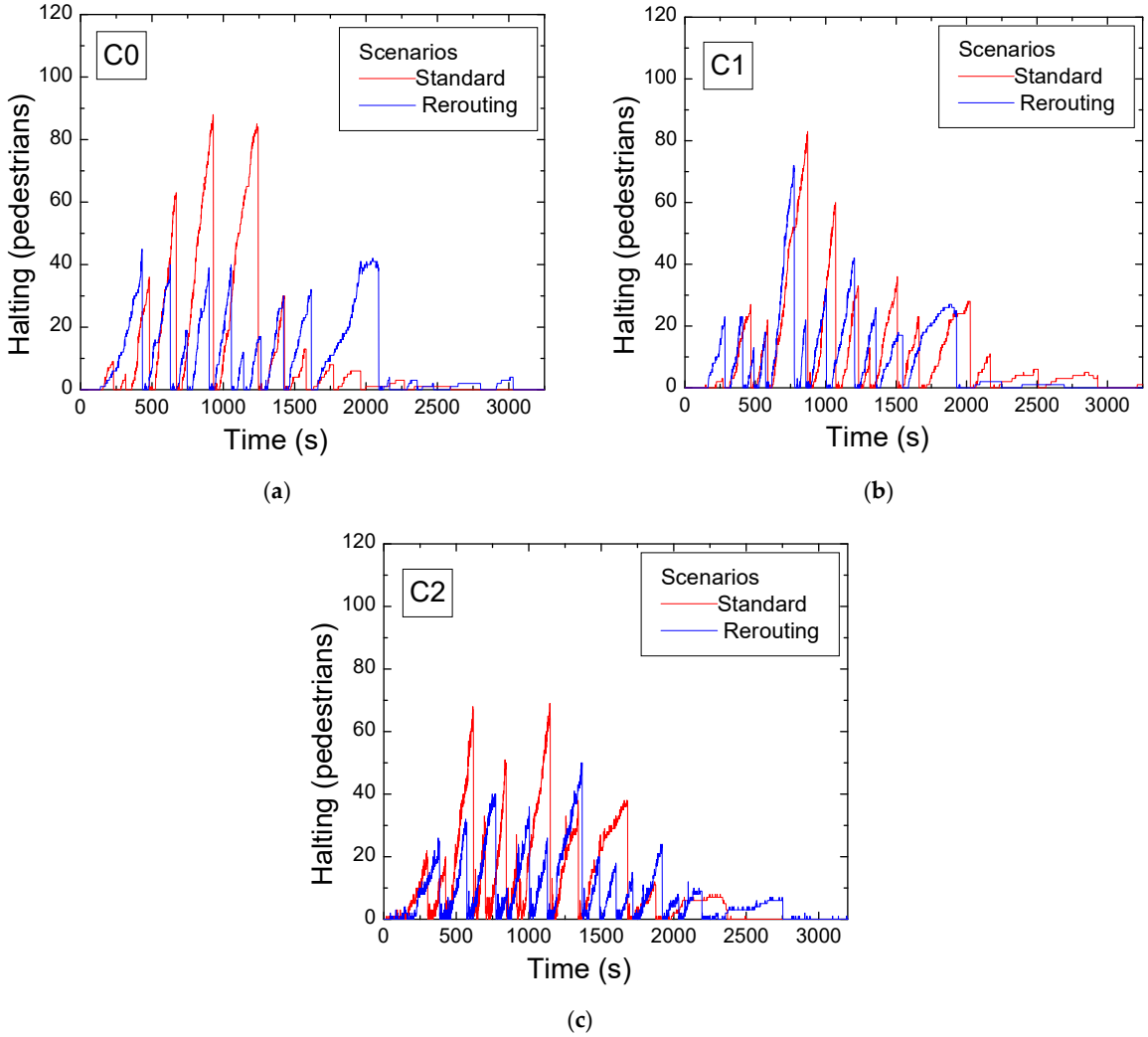


Figure 11. Comparison of trends over time for pedestrian halting sessions at intersections in standard versus rerouting scenario: (**a**) intersection C0, (**b**) intersection C1, and (**c**) intersection C2.

The difference between the scenarios is most pronounced at intersections C0 and C2, where the rerouting scenario more frequently utilizes the pedestrian phase, facilitating smoother pedestrian flow through these junctions. Intersection C1, however, sees a higher overall volume of pedestrians, resulting in a greater number of halted pedestrians compared to C0 and C2. Nevertheless, even at C1, the rerouting scenario shows fewer pedestrians waiting than the standard scenario.

While the rerouting scenario does not directly prioritize pedestrians, it has an indirect positive effect: with fewer vehicles in queues, the intersections run more efficiently, allowing more frequent activation of pedestrian phases.

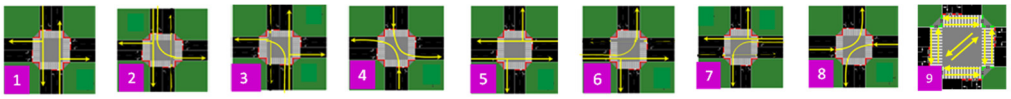
The pedestrian phase is crucial as it helps control vehicle flows and prevents excessive traffic on critical roads connecting intersections. In this sense, pedestrians benefit from the rerouting scenario. However, the overall reduction in pedestrian numbers is minimal because, even when a pedestrian phase is active, it does not always mean that many pedestrians will cross the intersection. This may be due to fewer pedestrians being present, smaller groups, or pedestrians not yet arriving in the waiting zones.

The peaks in the halting sessions are linked to crossing moments. The size of these peaks indicates the stress level and demonstrates a pedestrian's reaction to connected vehicles. Comparatively, smaller peaks are observed in the rerouting halting sessions, while higher peaks are observed in standard halting sessions. Compared to the dramatic pedestrian scenario, the rerouting scenario reduces the average pedestrian pressure by 25% between 500 and 1000 s. This demonstrates that rerouting effectively alleviates pressure during the critical period when both pedestrians and vehicles are at their peak waiting times, reducing the risk of pedestrian run overs. By minimizing potential conflicts and collisions, rerouting enhances the safety performance, emphasizing its importance for traffic management.

5.3. Global Agent Decisions

Table 3 presents the overall percentages of the green times for intersections C0, C1, and C2, over a training segment, for both scenarios. On the top, the nine possible phases are presented as a draft.

Table 3. Global percentages of green times for C0, C1, and C2, for both the standard and rerouting scenarios. On the top, the nine possible phases are presented as a draft for clarity.



Standard (% Green Time)				Rerouting (% Green Time)			
		C1	C2		C0	C1	C2
P1		15%	13%	P1	13%	17%	11%
P2	P2	10%	6%	P2	9%	5%	10%
P3	P3	5%	9%	P3	2%	6%	7%
P4	P4	2%	4%	P4	3%	1%	2%
P5	P5	30%	22%	P5	33%	32%	28%
P6	P6	7%	19%	P6	6%	12%	9%
P7	P7	5%	4%	P7	9%	9%	3%
P8	P8	10%	11%	P8	7%	4%	8%
P9	P9	30%	22%	P9	18%	14%	21%

The results show that the system prioritizes critical phases (P1, P5, and P9) in both scenarios, adapting its strategy to reduce waiting times and improve traffic flows. In the standard scenario, green times are predominantly allocated to the arterial direction (P5), decreasing sharply from 37% at C0 to 20% at C2. This approach creates controlled traffic flow bottlenecks across specific road sections or chains of sections, resulting in queue build-ups at junctions C0, C1, and C2, as shown in Figure 10a–c. In contrast, the rerouting scenario distributes green times more evenly, with P5 green times decreasing gradually from 33% to 28%, supporting a smoother and more balanced traffic flow across the network.

In both scenarios, Figure 12a,b depicts the temporal trends of all the active actions validated by the global agent at intersections C0, C1, and C2.

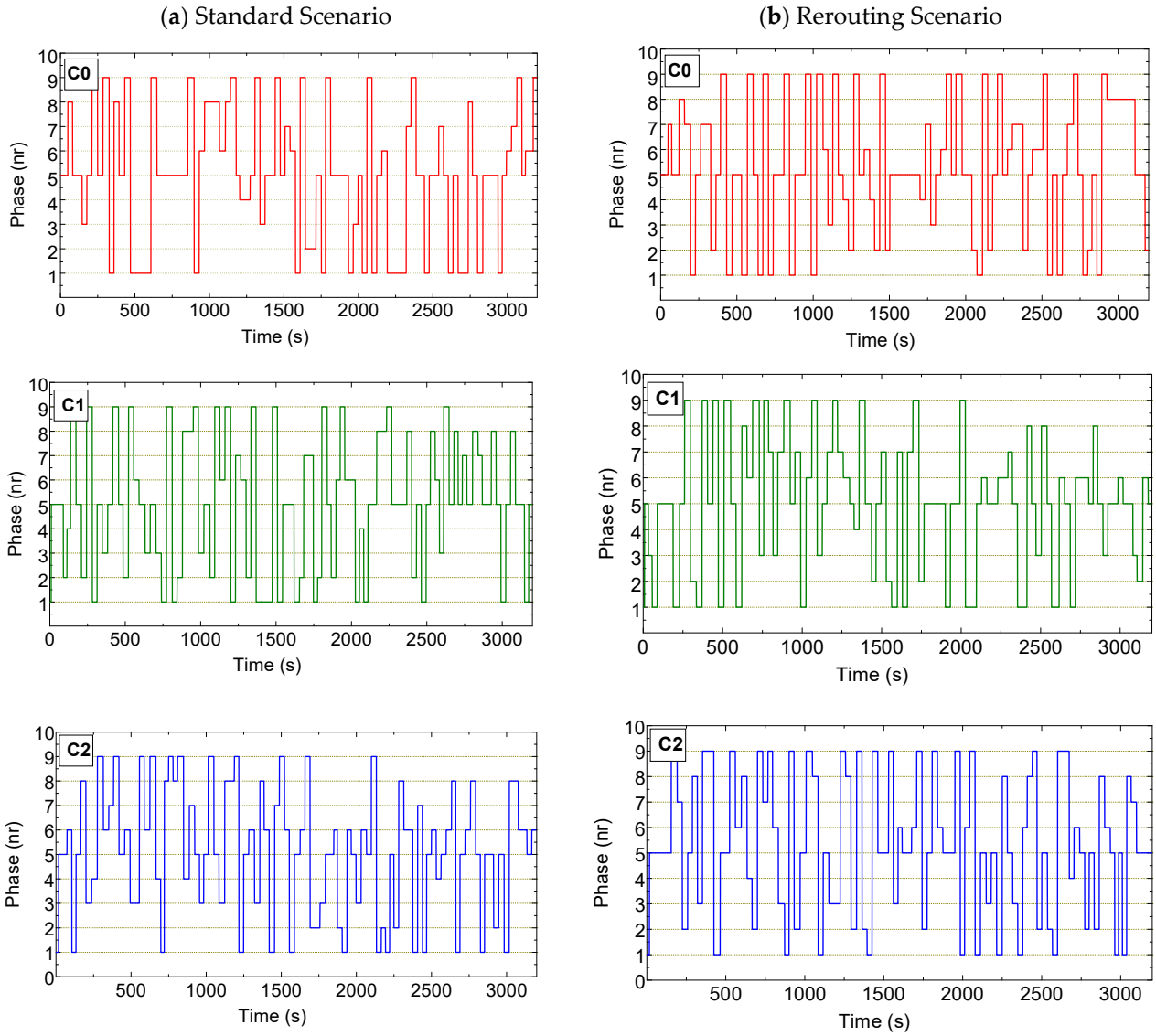


Figure 12. Comparison of trends over time for the active phases (actions) at intersections C0, C1 and C2. (**a**) Standard scenario. (**b**) Rerouting scenario.

The results show that the system diverges from a fixed phase sequence, characteristic of dynamic traffic control systems (Figure 3b), by continuously adapting to real-time traffic conditions. Crucially, pedestrian phases are only triggered upon pedestrian request, optimizing phase usage by prioritizing vehicular movements unless a pedestrian need arises.

When comparing both scenarios, the allocated green times and phase sequences across intersections vary over time. In the first 1500 s, green times for Phase 1 (N > S) and Phase 5 (E > W) at intersections C0 and C1 show marked differences from C2 in the standard scenario. In the rerouting scenario, green times for Phase 1 in C0 and C1 nearly double, while Phase 5's green time in C2 also increases compared to the standard configuration. Additionally, the rerouting scenario emphasizes phases 9 (pedestrians) and 8 (left turns), reflecting a more pronounced prioritization of pedestrian and turning movements to support smoother rerouting and congestion management.

Figure 13 illustrates the time-based trends for critical phases 5 and 1 across both scenarios at intersections C0, C1, and C2. An inset is included to clarify the active phase and scenario.

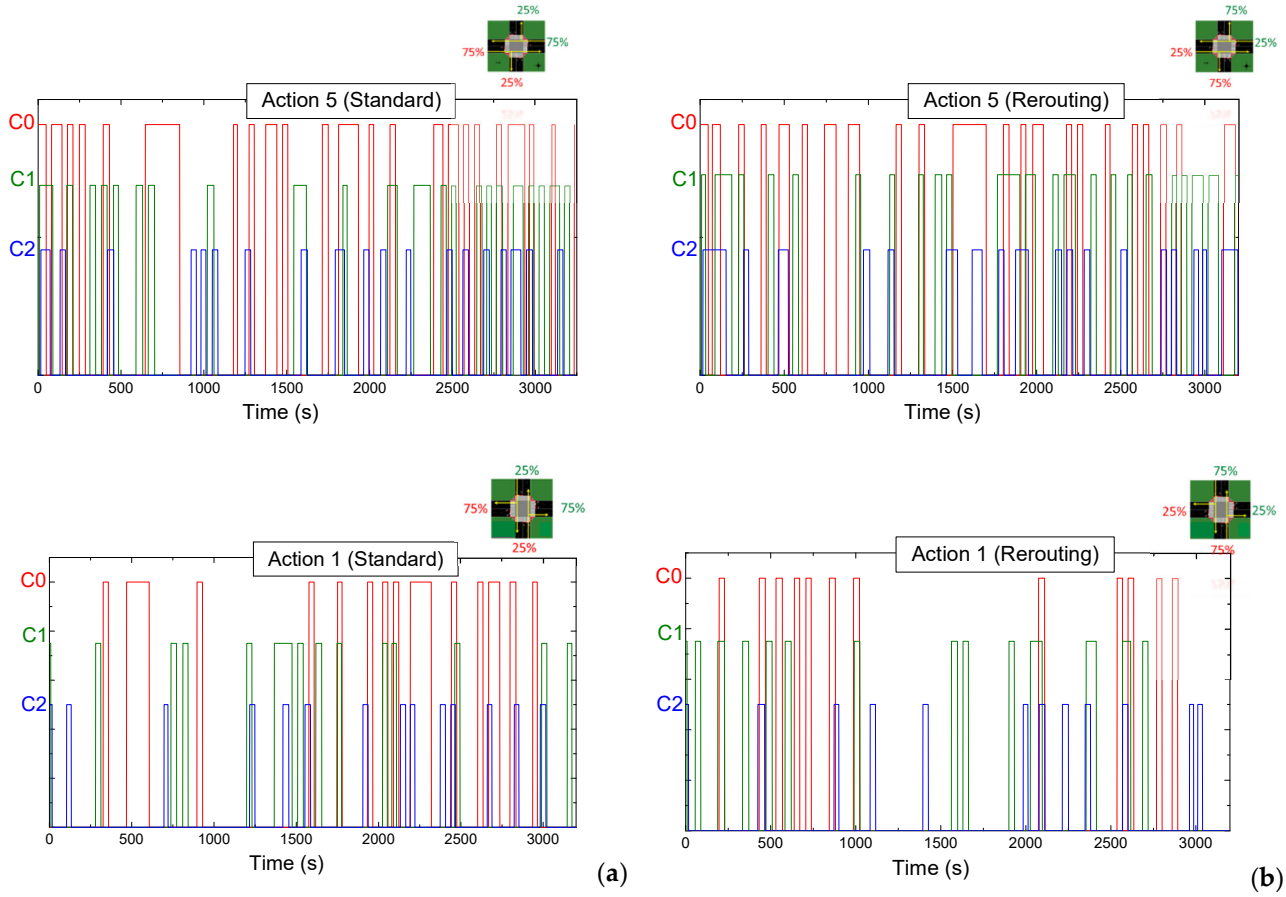
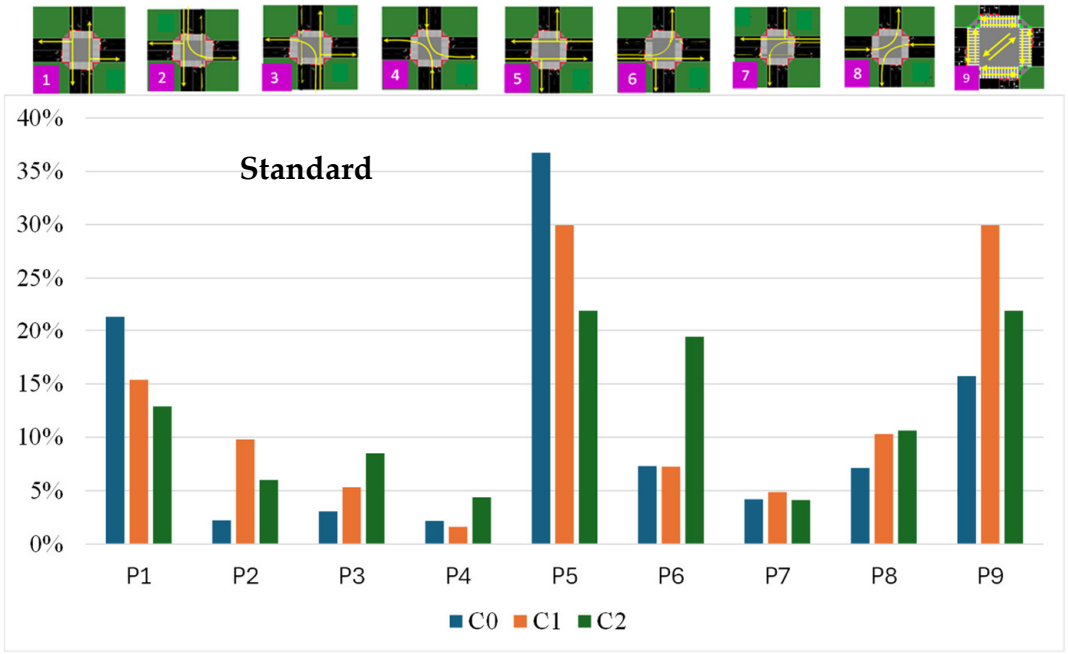


Figure 13. Time-based comparison of active phases 5 and 1 at intersections C0, C1, and C2: (**a**) standard scenario and (**b**) rerouting scenario. The phases and scenarios are shown in the insets.

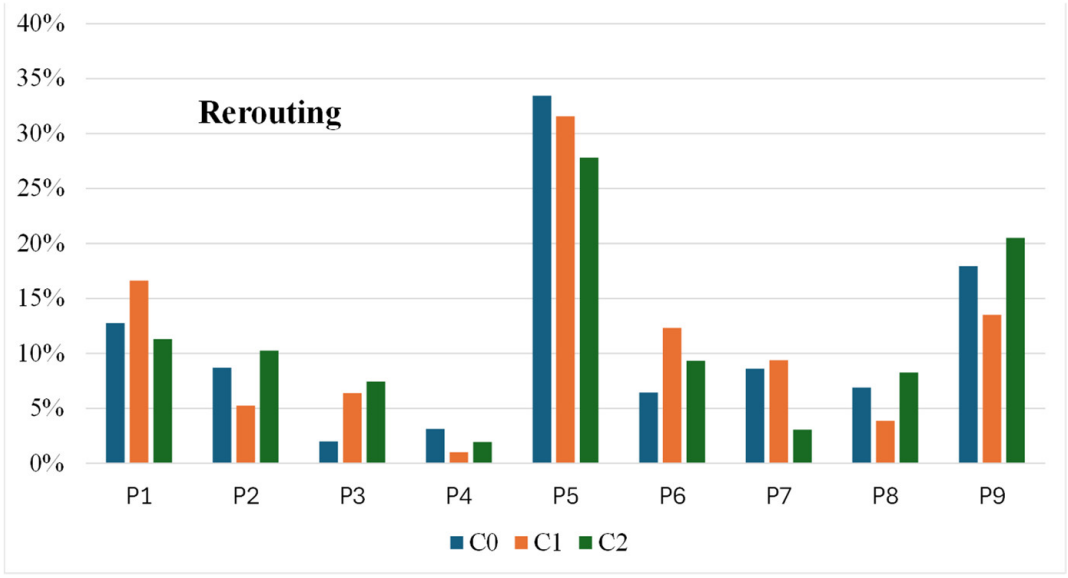
This adaptive approach optimally manages both pedestrian and vehicle flows. Initially, we began with a typical configuration of three junctions in a line: 75% of vehicles proceed straight ahead and 25% turn right or left in both directions. However, when arterial traffic demand surpasses system capacity, whether due to incidents or severe congestion, the rerouting scenario is activated. In this scenario, traffic light coordination is dynamically redesigned to adjust the vehicle distribution to 25% straight ahead and 75% turning. This shift reroutes a larger proportion of the vehicles onto alternative paths, significantly alleviating congestion along the main artery and enhancing the overall traffic flow. Through the rerouting mechanism, the system effectively reconfigures arterial coordination, reduces congestion, and strategically diverts traffic to less congested routes.

In Figure 14a, a comparison of the green time trends across all the active phases at intersections C0, C1, and C2 is displayed. Active phases are indicated at the top for clarity.

To reduce arterial congestion, the standard scenario increases the green times in P1 at C0 and C2, in P2 at C1, and in P6 at C2. Meanwhile, the rerouting scenario redistributes green time allocations to limit traffic on the arterial road. Pedestrian phase (P9) at C1 nearly triples in the standard scenario, enhancing safe-crossing opportunities. As a result, queues at C0, C1, and C2 are reduced, as shown in Figure 10a–c. The vehicles that are rerouted exit the environment via right turns, preventing additional pressure at downstream junctions.



(a)



(b)

Figure 14. Comparison of green time trends across all active phases at intersections C0, C1, and C2. Active phases are indicated at the top for clarity. (**a**) Standard scenario. (**b**) Rerouting scenario.

This adaptive system effectively manages both vehicle and pedestrian traffic. For vehicles, it minimizes critical queues, preventing congestion and relieving pressure at the intersections. For pedestrians, consistently activated crossing phases ensure safe, efficient movement through intersections.

Furthermore, the system activates the phases flexibly, without a fixed sequence, responding dynamically to real-time traffic patterns at each intersection. Overall, it accommodates 2600 vehicles and 2000 pedestrians, demonstrating strong adaptability to diverse traffic conditions.

Based on Figure 11, Table 3, references [22] and [24], and an evaluation of the key performance metrics, such as the percentage increase in the number of vehicles successfully passing through the intersections (Throughput Improvement), the percentage reduction in vehicle queue lengths (Queue Size Reduction), and the adaptability to dynamic traffic conditions, Table 4 provides comparative analysis of the traffic flow management methods. The proposed method is compared to VLC Dynamic Coordination (Fixed Phases [24]) with Intelligent Coordination ([22]) and Intelligent Coordination with Rerouting.

Table 4. Comparison of traffic flow performance metrics.

Coordination Method	Throughput Improvement (%)	Queue Size Reduction (%)	Adaptability to Traffic Variations
VLC Dynamic (Fixed-Phase Cycles) [24]	10%	15%	Low (static cycles unsuitable for variability)
VLC Intelligent Coordination [22]	30%	35%	Medium (dynamic adaptive phases and timings)
VLC Intelligent Coordination with Rerouting (this work)	63%	54%	High (dynamic adaptation and rerouting)

The results highlight the following:

- The *VLC Dynamic Coordination (Fixed-Phase Cycles)* approach demonstrates limited adaptability and minimal improvements, due to its reliance on predefined phase timings;
- The *Intelligent Coordination (Standard)* approach shows substantial improvements by integrating real-time traffic data to adjust signal phases dynamically;
- The *Intelligent Coordination with Rerouting* approach outperforms both methods, leveraging rerouting algorithms to enable vehicles to take optimized paths, significantly reducing congestion and delays.

This analysis underscores the effectiveness of the proposed method in improving traffic flow management and emphasizes its contribution to the field.

5.4. Impact of VLC and CV Integration: Challenges and Limitations

Integrating VLC real-time communication and rerouting into traffic management improves system responsiveness, efficiency, and safety for road users. Rerouting enables the dynamic synchronization of traffic signals across intersections, reducing stop-and-go driving, easing congestion, and ensuring smoother traffic flow. By sharing real-time data, the system anticipates queues and adjusts signal phases to minimize delays, benefiting both vehicles and pedestrians. It redirects traffic away from congested areas, prevents bottlenecks, and enhances pedestrian safety by coordinating crossings effectively.

However, rerouting effectiveness depends on fast, reliable data transmission. Communication delays can disrupt synchronization, causing inefficiencies and safety risks. Expanding the system to larger urban areas introduces scalability challenges, requiring robust computing infrastructure to handle the increasing data load. Many cities rely on outdated traffic control systems that may not fully support modern real-time technologies, making integration costly and complex.

To address these challenges, investments in resilient infrastructure, strategic upgrades, and efficient algorithms are essential for successful real-time, rerouting-based traffic management.

6. Conclusions and Future Trends

This paper presents an in-depth study of arterial intelligent traffic management systems across multiple intersections, highlighting the synergy between VLC technology and RL techniques to improve road safety and efficiency. Our work focuses on managing both vehicular and pedestrian traffic, with particular attention paid to optimizing the often-overlooked pedestrian phase. By integrating V-VLC messages with intelligent state representations, we developed dynamic traffic control models capable of efficiently managing traffic across interconnected intersections.

Using extensive simulations with MARL and the SUMO simulator, we demonstrated the effectiveness of our approach in optimizing traffic flow and reducing congestion. The analysis incorporated key factors, such as vehicle and pedestrian volumes, velocities, pedestrian clearance times, and waiting zone occupancy, essential elements for efficient traffic management.

The MARL model proved to be adaptable to various traffic scenarios, underscoring the importance of continuous learning in dynamic environments. Comparing the standard and rerouting scenarios at intersections C0, C1, and C2, showed that rerouting reduced the average traffic pressure by 66%, 50%, and 75%, respectively. These results highlight rerouting's ability to enhance arterial road efficiency and deliver practical benefits. Additionally, fewer stationary vehicles and shorter queues were observed when real-time traffic information from agents was utilized, effectively alleviating the pressure on intersections.

This redistribution of traffic improves decision-making for phase activation, enabling more frequent pedestrian phase activations when needed. Consequently, the overall flow of both vehicles and pedestrians improves, fostering a more balanced and efficient environment.

The primary focus of this work was to establish the baseline performance of the proposed system in standard conditions. This approach allowed us to evaluate its core functionality and establish a foundation for future research. Future research will explore the scalability and adaptability of our approaches across diverse urban settings and traffic and weather conditions. Testing in varied environments, including those with differing traffic volumes, road configurations, and pedestrian dynamics, will help refine and validate these methods. However, we recognize the importance of assessing the system's robustness in diverse environmental conditions, and we plan to address this in subsequent studies.

By advancing these techniques, we aim to contribute to the development of intelligent traffic management systems that address the complexities of urban mobility, while prioritizing safety and efficiency.

Author Contributions: Conceptualization, M.A.V.; formal analysis, M.V. (Manuela Vieira), G.G. and M.Ve. (Mario Vestias); investigation, G.G., M.A.V. and M.Ve. (Mario Vestias); methodology, M.V. (Manuela Vieira), M.A.V. and M.Ve. (Mario Vestias); software, G.G. and M.Ve. (Mario Vestias); validation, G.G., M.A.V., P.L. and P.V.; writing—original draft, M.V. (Manuela Vieira); writing—review and editing, M.V. (Manuela Vieira). All authors have read and agreed to the published version of the manuscript.

Funding: This research received support from FCT, Fundação para a Ciência e a Tecnologia, through the Research Unit CTS, Center of Technology and Systems, with the reference UIDB/00066/2020.

Data Availability Statement: All the data is in the manuscripts and figures.

Acknowledgments: The authors acknowledge CTS-ISEL and IPL.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

1. Siegel, J.E.; Erb, D.C.; Sarma, S.E. A survey of the connected vehicle landscape—Architectures, enabling technologies, applications, and development areas. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 2391–2406. [CrossRef]
2. Karagiannis, G.; Altintas, O.; Ekici, E.; Heijenk, G.; Jarupan, B.; Lin, K.; Weil, T. Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions. *IEEE Commun. Surv. Tutor.* **2011**, *13*, 584–616. [CrossRef]
3. Ge, H.; Song, Y.; Wu, C.; Ren, J.; Tan, G. Cooperative Deep Q-Learning With Q-Value Transfer for Multi-Intersection Signal Control. *IEEE Access* **2019**, *7*, 40797–40809. [CrossRef]
4. Vidali, A.; Crociani, L.; Vizzari, G.; Bandini, S. A Deep Reinforcement Learning Approach to Adaptive Traffic Lights Management. In Proceedings of the Workshop "From Objects to Agents" (WOA 2019), Parma, Italy, 26–28 June 2019; pp. 42–50.
5. O'Brien, D.; Le Minh, H.; Zeng, L.; Faulkner, G.; Lee, K.; Jung, D.; Oh, Y.; Won, E.T. Indoor Visible Light Communications: Challenges and prospects. In *Free-Space Laser Communications VIII, Proceedings of the Optical Engineering + Applications, San Diego, CA, USA, 10–14 August 2008*; SPIE: Bellingham, WA, USA, 2008; Volume 7091, pp. 60–68.
6. Parth, H.; Pathak, X.; Pengfei, H.; Prasant, M. Visible Light Communication, Networking and Sensing: Potential and Challenges. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 2047–2077.
7. Caputo, S.; Mucchi, L.; Cataliotti, F.; Seminara, M.; Nawaz, T.; Catani, J. Measurement-based VLC channel characterization for I2V communications in a real urban scenario. *Veh. Commun.* **2021**, *28*, 100305. [CrossRef]
8. Vieira, M.A.; Vieira, M.; Vieira, P.; Louro, P. Optical signal processing for a smart vehicle lighting system using a-SiCH technology. In *Optical Sensors 2017, Proceedings of the SPIE Optics + Optoelectronics, Prague, Czech Republic, 24–27 April 2017*; SPIE: Bellingham, WA, USA, 2017; Volume 10231, p. 10231.
9. Girao, P.S.; Alegria, F.; Viegas, J.M.; Lu, B.; Vieira, J. Wireless System for Traffic Control and Law Enforcement. In Proceedings of the 2006 IEEE International Conference on Industrial Technology, Mumbai, India, 15–17 December 2006; pp. 1768–1770. [CrossRef]
10. Chen, S.; Hu, J.; Shi, Y.; Peng, Y.; Fang, J.; Zhai, R.; Zhao, L. Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G. *IEEE Commun. Stand. Mag.* **2017**, *1*, 70–76. [CrossRef]
11. Oskarbski, J.; Guminska, L.; Miszewski, M.; Oskarbska, I. Analysis of Signalized Intersections in the Context of Pedestrian Traffic. *Transp. Res. Procedia* **2016**, *14*, 2138–2147. [CrossRef]
12. Pribyl, P.; Pribyl, M.; Lom Svitek, M. Modeling of smart cities based on ITS architecture. *IEEE Intell. Transp. Syst. Mag.* **2019**, *11*, 28–36. [CrossRef]
13. Miucic, R. *Connected Vehicles: Intelligent Transportation Systems*; Springer: Cham, Switzerland, 2019.
14. Kodi, J.H. Evaluating the Mobility and Safety Benefits of Adaptive Signal Control Technology (ASCT). Master's Thesis, UNF Graduate Theses and Dissertations, Jacksonville, FL, USA, 2019. Available online: https://digitalcommons.unf.edu/etd/930 (accessed on 10 November 2024).
15. Aleko, D.R.; Djahel, S. An Efficient Adaptive Traffic Light Control System for Urban Road Traffic Congestion Reduction in Smart Cities. *Information* **2020**, *11*, 119. [CrossRef]
16. Eldeeb, H.B.; Eso, E.; Jarchlo, E.A.; Zvanovec, S.; Uysal, M.; Ghassemlooy, Z.; Sathian, J. Vehicular VLC: A ray tracing study based on measured radiation patterns of commercial taillights. *IEEE Photonics Technol. Lett.* **2021**, *33*, 904–907. [CrossRef]
17. Bilal, J.M.; Jacob, D. Intelligent Traffic Control System. In Proceedings of the 2007 IEEE International Conference on Signal Processing and Communications, Dubai, United Arab Emirates, 24–27 November 2007; pp. 496–499. [CrossRef]
18. Yousefi, S.; Altman, E.; El-Azouzi, R.; Fathy, M. Analytical Model for Connectivity in Vehicular Ad Hoc Networks. *IEEE Trans. Veh. Technol.* **2008**, *57*, 3341–3356. [CrossRef]
19. Shen, W.-H.; Tsai, H.-M. Testing vehicle-to-vehicle visible light communications in real-world driving scenarios. In Proceedings of the 2017 IEEE Vehicular Networking Conference (VNC), Turin, Italy, 27–29 November 2017; pp. 187–194.
20. Liang, X.; Du, X.; Wang, G.; Han, Z. A Deep Reinforcement Learning Network for Traffic Light Cycle Control. *IEEE Trans. Veh. Technol.* **2019**, *68*, 1243–1253. [CrossRef]
21. Lopez, P.A.; Behrisch, M.; Bieker-Walz, L.; Erdmann, J.; Flötteröd, Y.-P.; Hilbrich, R.; Lücken, L.; Rummel, J.; Wagner, P.; Wiessner, E. Microscopic Traffic Simulation using SUMO. In Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2575–2582.
22. Vieira, M.; Vieira, M.A.; Galvão, G.; Louro, P.; Véstias, M.; Vieira, P. Enhancing Urban Intersection Efficiency: Utilizing Visible Light Communication and Learning-Driven Control for Improved Traffic Signal Performance. *Vehicles* **2024**, *6*, 666–692. [CrossRef]
23. Yousefpour, A.; Fung, C.; Nguyen, T.; Kadiyala, K.; Jalali, F.; Niakanlahiji, A.; Kong, J.; Jue, J.P. All one needs to know about fog computing and related edge computing paradigms: A complete survey. *J. Syst. Archit.* **2019**, *98*, 289–330. [CrossRef]
24. Vieira, M.A.; Vieira, M.; Vieira, P.; Fernandes, R.; Louro, P. Dynamic vehicular visible light communication for traffic management. In *Next Generation Optical Communication: Components, Sub-Systems, and Systems XII*; Li, G., Nakajima, K., Srivastava, A.K., Eds.; SPIE: Bellingham, WA, USA, 2023; p. 124290.
25. Vieira, M.A.; Vieira, M.; Louro, P.; Vieira, P.; Fantoni, A. Vehicular Visible Light Communication for Intersection Management. *Signals* **2023**, *4*, 457–477. [CrossRef]
26. Vieira, M.A.; Vieira, M.; Louro, P.; Vieira, P. Cooperative vehicular communication systems based on visible light communication. *Opt. Eng.* **2018**, *57*, 076101. [CrossRef]

27. Tang, F.; Kawamoto, Y.; Kato, N.; Liu, J. Future intelligent and secure vehicular network toward 6g: Machine-learning approaches. *Proc. IEEE* **2019**, *108*, 292–307. [CrossRef]
28. Luong, N.C.; Hoang, D.T.; Gong, S.; Niyato, D.; Wang, P.; Liang, Y.-C.; Kim, D.I. Applications of deep reinforcement learning in com munications and networking: A survey. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 3133–3174. [CrossRef]
29. Ye, H.; Li, G.Y.; Juang, B.-H.F. Deep reinforcement learning based resource allocation for v2v communications. *IEEE Trans. Veh. Technol.* **2019**, *68*, 3163–3173. [CrossRef]