RESEARCH ARTICLE

A Decentralised Multi-Agent DRL-based Approach for Pedestrian and Vehicle Traffic Signals Controlling Systems Optimisation

Mohammed Anis Oukebdane[1,*]

[1] Department of Electronics and Communication Engineering, Yildiz Technical University, Istanbul, Turkey

Abstract

Urban traffic congestion is a major issue that negatively affects mobility efficiency, environmental sustainability and road safety. Many recent methods for controlling traffic signals have used methods based on deep reinforcement learning (DRL) and provided positive results. However, it focused primarily on vehicle flow and have not taken into account pedestrian dynamics due to inherent difficulty related to accurately sensing all pedestrians. As a result of these limitations, recent advances in sixth-generation (6G) localisation technology will provide new opportunities to provide precise, low-latency tracking of pedestrians at signalized intersections, allowing for improved control of pedestrian movements in urban areas. The model proposed in this paper named DRL-based pedestrian-vehicle traffic signal management (DRL-PVTSM), provides a solution to this need by providing a decentralized multi-agent DRL approach that jointly optimizes both vehicle and pedestrian movements at each intersection using independent agents each controlled by deep Q-network (DQN). The agents are provided with a pressure-based reward for optimizing vehicle and pedestrian queue densities and have created safety-penalizing rewards based on pressure from pedestrians that are waiting for the lights to change. The DRL-PVTSM framework has been designed in accordance with the principles of scalability, robustness and real-time applicability to large multi-intersection urban traffic networks. This work demonstrates in extensive simulations performed in SUMO software on multiple network traffic topologies of grid and random layouts that the DRL-PVTSM model provides statistically significant improvements in pedestrian waiting time, vehicle travel delay, and decreases in congestion mitigation and intersection-level safety indicators, thus confirming that decentralized DRL with future 6G will provide a viable method for optimizing the joint operation of pedestrian and vehicle traffic signal systems.

Keywords: scalable adaptive traffic signal control, decentralized multi-agent deep reinforcement learning, pedestrian-vehicle coordination, traffic signal optimization, 6G.

Submitted: 31 December 2025
Accepted: 09 February 2026
Published: 21 March 2026

Vol. 2, **No.** 1, 2026.
10.62762/TMWI.2025.878487

***Corresponding author:**
✉ Mohammed Anis Oukebdane
mohammed.oukebdane@std.yildiz.edu.tr

Citation
Oukebdane, M. A. (2026). A Decentralised Multi-Agent DRL-based Approach for Pedestrian and Vehicle Traffic Signals Controlling Systems Optimisation. *ICCK Transactions on Mobile and Wireless Intelligence*, 2(1), 31–43.

© 2026 ICCK (Institute of Central Computation and Knowledge)

1 Introduction

Traffic congestion is one of the most pressing urban problems in the world today and leads to significant economic, environmental and safety impacts [1]. A recent study conducted by the texas transportation institute (TTI) estimated that traffic congestion costs the urban areas of the united states approximately $100 billion per year due to wasted fuel, lost productivity, and increased risk of accidents (see Figure 1) [2]. Therefore, there has been a growing interest in optimising the operation of traffic through efficient infrastructure use. Adaptive traffic signal control is likely one of the best solutions available today to alleviate the issue of traffic congestion while increasing the possibilities for movement. These new traffic control systems allow for the length of time for which a traffic signal is green to be adjusted dynamically as real-time changes to traffic conditions occur.

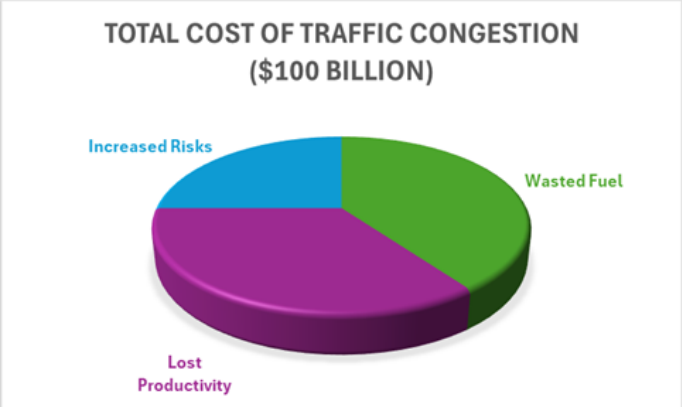


Figure 1. Total cost of traffic congestion ($100 billion).

Recent years have seen substantial advances in the utilisation of reinforcement learning (RL) technologies and methodologies for the establishment of adaptive traffic signal control (ATSC) systems based upon artificial intelligence (AI). Unfortunately, most current applications of RL in the transportation domain tend to concentrate primarily on optimising vehicle movement patterns: they frequently do not consider how to coordinate pedestrians' movements with those of the vehicles, due mainly to the challenging nature of monitoring and tracking pedestrian activity at intersections in a dependable manner [3, 4]. Vision systems routinely employed for the purpose of monitoring and tracking pedestrians on the road have generally proven ineffective in the context of outdoor environments that are subject to dynamic obscurations (e.g., other vehicles) as well as variations in environmental conditions such as light and weather [5]. However, as wireless communication networks transition from fifth-generation (5G) to sixth-generation (6G) technologies, they will provide a means by which to enhance pedestrian monitoring and tracking accuracy (in real time) at traffic signalised intersections [6, 7]. In addition to the advantages associated with 6G networks, the combination of highly detailed pedestrian traffic data with vehicular patterns presents the first opportunity to jointly optimise the interaction between pedestrians and vehicles at traffic signalised intersections, thus allowing for improved transportation system efficiencies, safety, and mobility throughout many urban centres.

In this paper, an adaptive traffic signal control model is being developed for cities with many intersections that have a mix of pedestrian and vehicle traffic. To create a real-time situational awareness of these areas, we used the SUMO simulator, which allow for the flexible and timely signalling of all intersections within a simulated city. In addition to optimising for vehicular efficiency and pedestrian movement, this method increases its ability to address some of the deficiencies that exist in current fixed-time vehicle control strategies, while also increasing its capacity and effectiveness when expanded to include much larger deployments. This study provides contributions to the field of TCS as follow:

- A multi-agent deep reinforcement learning (DRL)-ATSC architecture was developed for multiple intersection networks that considers mixed pedestrian-vehicle traffic, enabling coordination of signal decisions across multiple intersections rather than having vehicle-centric designs only.
- Utilisation of SUMO with specific configurations to provide fine-grained, low-latency information on pedestrian states; this information is sent to the real-time proposed system, and the model reacts accordingly in as short time as possible.
- Contrary to traditional and current solutions, the proposed model establishes a unified control objective to jointly optimise traffic efficiency and pedestrian safety, thus resolving the inherent imbalance present in conventional fixed-time and vehicle-only signal plans.
- A scalable coordination mechanism to coordinate multiple intersection networks in large urban areas, while still maintaining manageable communication and computational complexity.

- Comprehensive performance evaluation which shows significant improvement over fixed-time and vehicle-only adaptive control systems in regards to delay, throughput, and pedestrian service quality.

This paper is structured as follows: next, in Section 2 we present the current studies in this field and highlight its limitations. Followed by Section 3 that presents the methodology of this work. Section 4 showcases the proposed model's design. Next, in Section 5, the detailed steps used to perform this work are cleared. Following that, we show the obtained results in Section 6 and we discuss the future directions in Section 7. Finally, we end the paper with a conclusion as in Section 8

2 State of the art

The growing interest in adaptive TSCs as a viable solution for resolving urban congestion has prompted further advances in new research and innovations towards such systems. Recent work has introduced multiple models of control policy solution creation using manual analytical methods to model traffic flows & control policies [8]. Limitations from such methods include the need for specifying an accurate set of model definitions for every intersection that a signal will operate at so to maximize scalability and robustness if an intersection's similar traffic flow patterns are not being followed. The emergence of RL methods is providing data-driven alternatives that are enabling the creation of flexible and scalable signal operation policies that support the optimization of signals based on the use of empirical traffic data. Recent value-based signal operation approaches utilizing Q-Learning have demonstrated the capability to support the development of optimal traffic signals [9], while policy network approaches developed through the use of multi-agent reinforcement learning (MARL) have demonstrated success in developing optimal traffic signal policies for larger-scale networks. The ability to include and prioritize pedestrians within these earlier versions of traffic signal management systems continues to be significantly underdeveloped. Our deep reinforcement learning pedestrian-vehicle traffic signal management (DRL-PVTSM) approach contributes to these bodies of work by integrating precise pedestrian data unlocked by emerging 6G localisation capabilities with MARL formulations for scalable mixed traffic coordination.

There has been extensive research done regarding methods to optimize the flow of vehicular traffic through various types of intersections, optimizing the flow of vehicular traffic may help alleviate traffic congestion caused by urbanization, improve traffic accident rates, and improve the overall efficiency of urban transportation systems. A recent trend in adaptive traffic signal control systems (ATSC) is the use of real-time adaptive traffic signals to automatically adjust traffic signals based on real-time traffic data. Before the creation of ATSC, early research conducted on traffic signal control primarily employed analytical models that required researchers to develop accurate models of every individual intersection [11]. Although these types of models provided researchers with valuable insights into how to control intersections, they were limited in that they lacked scalability and flexibility when the actual flow of vehicular traffic at any particular intersection differed from the assumptions that were made during the model's development [12–14]. The introduction of machine learning (ML), such as with the use of RL, has revolutionized the ATSC system. Machine learning (ML)-based studies, particularly RL, have revolutionized ATSC. Researchers have developed value-based methods (Q-Learning policy gradient) that utilize real-world traffic data to help optimize traffic light timing [15]. For example, research has shown that RL can be an effective way of controlling motor vehicle traffic, but there is a huge gap in works incorporating congestion due to pedestrian traffic ignorance [16, 17].

ATSC technology is improved by new AI Technologies and the application of DRL will help improve traffic signal timing [15]. This technology enables the optimization of signal timing by using multiple distributed multi-agent systems [13, 19, 21]. By giving agents control of intersections (i.e. traffic signals), each agent can learn how to react autonomously to changing conditions in their local area [19].

Detecting and tracking the movement of pedestrians has been one of the greatest challenges associated with optimising traffic signals. Vision based technologies have been hampered by changing dynamic occlusions, weather and lighting conditions [22, 23]. However, the upcoming 6G networks provides enhanced localisation and tracking capability allowing for accurate and timely collection of pedestrian traffic data [24].

By using data from both vehicle and pedestrian traffic, it is possible to optimise overall mixed traffic. In previous studies, it has been shown that using pedestrian data along with vehicle data results in a

more efficient, safe, and mobile traffic management system than using either type of data alone [25]. Using pedestrian and vehicle data together allows developers to create better solutions that provide a balance between optimising traffic flow while providing for the safety of pedestrians at intersections [26]. Researchers has found that distributed multi-agent systems can be used to manage the control of traffic lights to improve overall traffic efficiency. Agent collaboration and competition interactions improve the ability to manage traffic flow [14]. The use of distributed multi-agent DRL will allow for the independent modification of the signal lights at each intersection while simultaneously contributing to the overall improvement of the traffic network [27].This study discusses the innovative use of simulation to evaluate the effectiveness of ATSC strategies and introduces future technology trends to enhance the ATSC. In most valuable studies a specific software for simulation was used which is called SUMO (Simulation of Urban MObility) to test and refine traffic signal control strategies incorporating multiple methods of evaluating ATSC [28].

The rapid advancement of new technologies have allowed for the development of many types of technologies that will benefit ATSC, including edge computing (EC). Future technologies, including 6G, can provide new opportunities to improve ATSC, such as real-time monitoring of traffic conditions, ATSC and improved pedestrian safety as a result of unmanned aerial vehicle (UAV) usage [29, 30]. Using unmanned aerial vehicles (UAV) can provide a bird's eye view of traffic patterns with the added benefit of being able to capture images and collect real-time information about how traffic is currently flowing. The integration of UAVs into ATSC would create improved accuracy and efficiency when compared with traditional methods of collecting and reporting on traffic data [31].

Another area of research that has gained significant interest is the development and implementation of intelligent transportation systems (ITS) using vehicle-to-infrastructure technologies. ITS provide improved traffic signal management and minimize the potential for congestion through the linking of vehicles with their surrounding infrastructure [32, 33].

By adding more internet of things (IoT) devices, the ATSC is able to have many different sources of information to improve how the traffic moves on the roadway. Sensors and other IoT devices will provide a way to capture data on vehicles, signals, and roadways as they happen. This data can then be sent back to the ATSC for real-time response and adaptation which allows it to improve the flow of traffic more quickly [34, 35].

The continued integration of future technologies such as intelligent UAV networks for traffic monitoring, smart crosswalks (i.e., integrating sensors to detect pedestrians in need of assistance), and multi-modal transportation systems (i.e., combining several modes of transportation into one) creates the ability to improve the efficiency, safety, and sustainability of urban transportation systems. The importance of big data analytics, or the analysis and aggregation of large amounts of historical traffic data, into traffic control systems has also increased. By analyzing large amounts of traffic data, many patterns and trends can be identified that can help develop more accurate TMSs [36]. The concept of blockchain was developed to provide secure and transparent means for transferring and exchanging data between TMSs, resulting in providing a higher quality of traffic data, increasing the level of trust in how the various TSCs share traffic data with one another and, ultimately, improving ATSC' ability to communicate with each other [15].

The future development of a complete ATMS that will utilize innovative wireless broadband communication systems and ML to better manage the interaction between people and vehicles in cities is imminent. It will provide new tools that will enhance the safety and performance of urban transportation networks through enhanced knowledge of pedestrian traffic flows. A fully ATMS has the potential to utilize both accurate pedestrian monitoring and decentralized MARL algorithms to optimize the flow of mixed cars and pedestrians through the city's streets and roads. Improved safety and efficiency of urban transport networks will enable a more uniform methodology of transportation among all modes of transportation, including mechanical (motor vehicle), human (pedestrian) and non-mechanical (bicycle) modes.

3 Methodology

3.1 Problem formulation

The multi-intersection signal control problem has been modelled as a Markov decision process (MDP) so that decentralised RL techniques can be applied [37]. The components of the MDP are defined below:

- States (s): a state describes how an intersection is currently performing. A state will include as a minimum: vehicle density, pedestrian queues,

and signal phase (green, yellow, red).

- Actions (a): an action is a signal phase (change or maintain) and the order of phases for a signal is cyclical, thus allowing for safe pedestrian crossings.
- Transition (T): the combination of actions and states will determine what happens to traffic during transitions. All vehicle density information is based upon how the system has operated in the past.
- Rewards (R): trade-offs between efficiency and safety; thus, a mixture of pressures from vehicles, pedestrians, and jaywalking penalties will be used.

The suggested control issue for pedestrian and vehicle traffic signals is represented as a Markov Decision Process (MDP) characterised by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$. At time step t, intersection agent i sees the state $s_i(t) \in \mathcal{S}$, which includes the lengths of the queues of vehicles on all incoming lanes, the lengths of the queues of pedestrians on each crossing, the currently active signal phase, and the time that has passed since that phase started. The action space $\mathcal{A}$ consists of distinct choices regarding signal phases, including phase switching and phase extension. The SUMO simulator changes the flow of traffic to control the transition dynamics $\mathcal{P}$. To specifically quantify congestion imbalance, vehicular pressure is characterised as the disparity between upstream and downstream queue lengths, whereas pedestrian pressure is quantified as the total waiting demand at crosswalks. The reward function combines both parts and is designed to reduce the combined pressure on vehicles and pedestrians while punishing unsafe interactions between pedestrians and vehicles. This guides the learning process toward safe and efficient signal control.

The goal here is to make rules that are specific to each intersection and get the most cumulative discounted rewards while also thinking about how they will affect traffic patterns in the long run.

3.2 Decentralised multi-agent deep Q-network (DQN)s

Using decentralised multi-agent DQNs, we designed our DRL-PVTSM scheme with the following characteristics [10]:

- Encoding States: The state encoding includes the amount of vehicles in an intersection, the queues of pedestrians, the present phase of the traffic signal (G, Y, R) relative to how quickly a pedestrian walks, and how long a pedestrian has to cross the intersection when the light is red.
- Actions: The actions include changing or continuing with the current phase of the traffic signal and allow for the dynamic expansion of green lights based on pedestrian vehicle counts.
- Phases: The traffic signal has a patterned sequence of regular and all-red phases, which allows for safe switching between the two.
- Rewards: To get more work done, you have to put pressure on people, and to keep pedestrians safe, you have to punish people for jaywalking.
- Learning: Massive user-based simulation for rapid retrieval of DQN optimisation data.

3.3 Implementation details

Using the SUMO simulation tool, we will extensively put the DRL-PVTSM model under a variety of baseline systems. The SUMO simulator realistically replicates the heterogeneous dynamics of vehicles and the complex behaviour of pedestrians in an urban environment and provides complete control over the design of the roadway network (e.g., geometries of streets and intersections) and the operation of traffic signals, the vehicle and pedestrian movement flows, and other elements of the urban traffic environment [13]. The SUMO simulator will be configured to represent a grid-like topological structure consisting of 1-by-4 grid layouts of 16 intersections, with each intersection connected by a 500-metre-long street segment, and will therefore provide an opportunity to fairly compare all simulation runs of the other systems, as all of the systems will utilise the same traffic layouts and traffic generation algorithms. The simulated traffic flows will consist of varied mixes of vehicles (e.g., cars, trucks, etc.) and various pedestrian types (e.g., slow-moving, fast-moving, etc.) and will be randomly assigned to origin-destination (OD) pairs of the various traffic flows. The pedestrian walking rates and pedestrian crossing times will be calibrated based on the pedestrian behaviour models as found in [16].

4 DRL-PVTSM design

To create the DRL-PVTSM system, we combined decentralised MARL with a model that coordinates pedestrian and vehicular traffic flows with an adaptive signal control system (SUMO). In this model, each

intersection is controlled by its own DQN agent and is able to make decisions based on local traffic conditions. Every DQN agent can learn from its own trial-and-error experience of traffic signals. Our learning method differs from previous vehicle-centric control methods by incorporating both pedestrian behaviour and safety needs as part of the decision-making process. To allow for this mixed demand of traffic, we combine both the amount of vehicles waiting in vehicle queues (in the form of densities) and the number of pedestrians waiting in pedestrian queues (in the form of densities) to form our state representation. We create a cyclical phase encoding scheme that can dynamically extend green signal times based on whether or not pedestrian crossings are still active. Additionally, we have developed a reward function that balances the efficiency objectives of traffic signals with the pedestrian priority objectives. We have also incorporated some new techniques to accelerate the learning process for the DQN agents, including a method for assigning priorities for experience replay, the use of ϵ-greedy exploration, and parallel simulation of DQN agents. The design of this ATSC is also intended to be practical, safe, and scalable based on actual constraints that exist at intersections in urban areas.

Algorithm 1: DRL-PVTSM training procedure.

Initialise SUMO environment
for each intersection agent i **do**
  Initialise DQN parameters θ_i and target network θ_i^-
  Initialise replay buffer $\mathcal{D}_i$
end for
for each episode **do**
  Reset environment
  for each time step **do**
    for each agent i **do**
      Observe state s_i
      Select action a_i using ϵ-greedy policy
    end for
    Execute joint actions in SUMO
    Observe rewards r_i and next states s_i'
    for each agent i **do**
      Store (s_i, a_i, r_i, s_i') in $\mathcal{D}_i$
      Sample mini-batch and update θ_i
    end for
  end for
end for

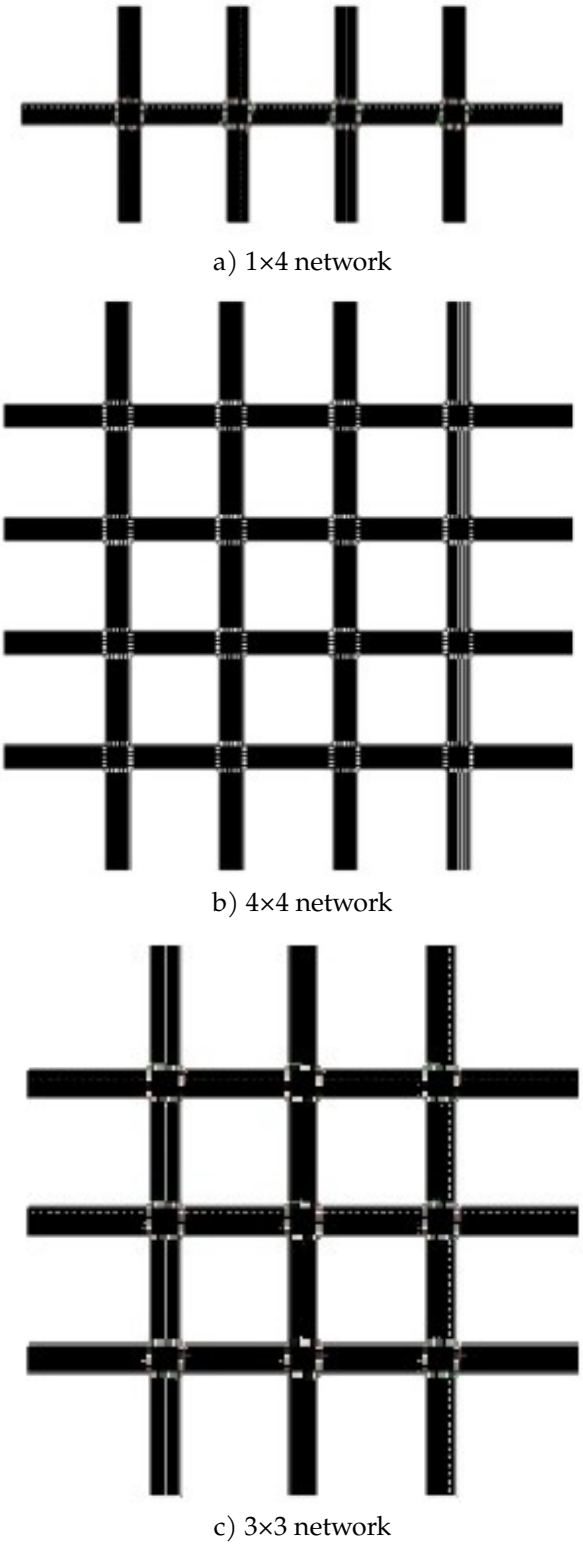
The DRL-PVTSM training method shown in Algorithm 1 employs a decentralised independent learning mode where an autonomous DQN agent controls each signalised intersection and interacts with the traffic environment through a simulation platform known as SUMO. In each episode of training, the simulation environment is reset and it will advance or develop in discrete time periods, and at each of those time periods, each agent will observe its local environment which includes both vehicular and pedestrian queue information and the active signal phase. Each agent will then make a decision about a control action by using an epsilon-greedy policy based on its current Q-network. After the execution of the control actions made by all agents in a synchronous manner in the SUMO environment, the agents will receive updates to their traffic states and will each receive a locally computed reward based on the performance of a particular control action according to both pressure and safety factors. Each agent will store the transition tuple in its replay buffer and will periodically sample the replay buffer to update its network parameters by using stochastic gradient descent, while a separate target network will be employed to enhance learning stability. While agents are independently learning and are not requiring mutual communication, the joint behaviours of all agents in the SUMO traffic simulation are inherently coordinated because all agents are connected to the same traffic flow systems, and thus, the actions taken by one agent may have an effect on the downstream states that are received by its adjacent agents. By enabling the use of this decentralised strategy for training an agent, this method allows for decentralised deployment and maintains real-time performance and stable learning, in a multi-intersection urban network.

In our analysis of the simulation of the DRL-PVTSM model, we focused on three different types of configurations: 4 intersections by 4 intersections (4x4), 1 intersection by 4 intersections (1x4), and 3 intersections by 3 intersections (3x3), as shown in Figure 2. These different types of configurations are represented as black dots on the urban road grid. For this paper, we primarily use the 1x4 configuration (Figure 2(a)) as our test case for analysis. The design for this type of road grid consists of 16 sequential intersections that are aligned in a vertical arrangement of 4 streets, where each street is spaced apart from one another by a horizontal distance of 500 metres. In the simulation for all experiments presented in this article, traffic flows with a mixture of vehicular and pedestrian

traffic are generated at random from all parts of the network to each of the 16 intersections, with the origins and destinations of all vehicle and pedestrian traffic being random throughout the network. Each intersection has been assigned a separate agent that operates as an autonomous agent and is trained to control the timing of traffic signals at all intersections based solely on information collected at the individual intersection throughout the entire duration of the experiment. By using a simplified design for the 1x4 grid to benchmark the ability to handle the directional throughput fluctuations and the need for coordination of traffic signal timings between the intersections of the 4x4 design prior to moving on to testing of the more complex designs, an overall improvement in the ability to effectively manage the increased densities and overall lengths of the road networks as per the original intent of the research study will be provided. The use of a modular road grid and the opportunity to gradually increase the density of traffic and network dimensions, as well as to continuously evaluate the overall capability of each agent through the training and evaluation phases, makes it possible for us to assess the effectiveness of this model and to subsequently make improvements to the overall design of the DRL-PVTSM model.

5 Implementation

This investigation is inspired by the research of Xu et al. [16]. To ensure consistency and comparability with the original study, we replicated the topological road networks described in their work when constructing our experimental setup. In addition, to extend the scope of the analysis and evaluate the robustness of the proposed parameters, a third road network topology (shown in Figure 3) was introduced, consisting of 147 nodes connected by 434 edges. Following the methodology described in [16], each road segment was set to a length of 500 meters, with sidewalks added at a width of 3 meters. The vehicular traffic component of the network was configured with a maximum speed of 45 km/h. A key distinction in this study lies in the pedestrian parameters, where the maximum walking speed was set to 3 m/s rather than matching the vehicular speed of 45 km/h. By differentiating maximum walking and driving speeds, this study examines the interactions between vehicular traffic and pedestrians, and how these interactions influence movement dynamics within the road network over time. These parameter adjustments broaden the analytical scope and contribute to a more comprehensive understanding of mobility behavior in simulated urban environments.



a) 1×4 network

b) 4×4 network

c) 3×3 network

Figure 2. Road network topologies.

5.1 Generating random roads and trips

To support the fair and unbiased execution of our study, we have created a custom-built Python application to randomly generate 224 unique combinations of trips

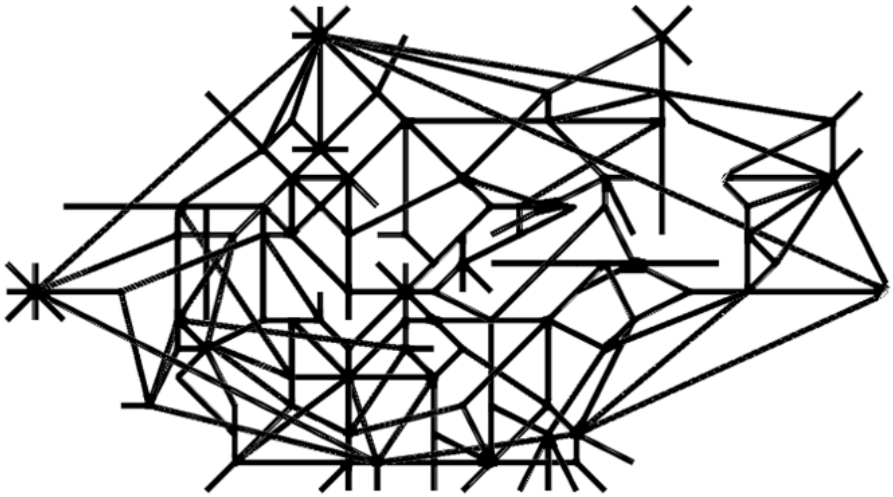


Figure 3. Random road network topology.

Table 1. Results of the experiments.

Metric	1 × 4		3 × 3		4 × 4		Random	
	P	V	P	V	P	V	P	V
Route length	1558.26	1821.50	1425.82	2113.40	1717.73	2496.35	1196.40	1886.95
Speed		13.18		14.22		15.52		5.27
Duration	1298.28	163.10	1183.87	163.43	1433.01	171.21	1173.97	400.80
Waiting time		54.35		45.56		44.35		264.81
Time loss	177.28	119.88	160.77	114.97	199.20	112.19	310.32	352.03
Depart delay		32.83		17.32		9.13		1.21

and roadways (scenarios) for both pedestrians and vehicles. We have also developed a data converter that exports all the data generated by our application into a format that can be used with SUMO. The combination of our Python application, a well-designed data extractor, and SUMO creates an opportunity for us to provide statistically meaningful output. The followings provide a detailed look at the quantitative data collected and analysed from our project. As part of this, we will present these results in Table 1 for the reader's reference.

The primary intent in developing the custom Python application was to provide a source of statistically reliable information that would allow us to create 224 distinct combinations of trips and roadways that were accurately representative of typical pedestrian and vehicular interactions. The fact that the generated combinations of trips and roadways were randomly selected from an entire set creates an accurate representation of a full complement of trip and road interaction data for use within the context of our research. As previously stated, our use of these tools and technologies, combined with the fact that SUMO has an established reputation as a leading microscopic traffic simulation platform, lends credibility to the methodology of our research and the process used to collect and analyse the data.

All statistical outputs and results obtained from the combination of SUMO and our Python application are presented in Table 1 for review and analysis. This table allows all interested parties to assess the simulation results and find useful insights from our research.

5.2 Network-level trip generation and SUMO integration

The process used to generate the findings in Table 1 outlined in this subsection is defined in an implementation-agnostic manner using algorithmic abstractions and mathematical notation to facilitate reproducibility, scalability, and compliance:

5.2.1 Network pre-processing and auxiliary operations

Let the traffic network be modelled as a directed graph

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}), \tag{1}$$

where $\mathcal{V}$ denotes the set of junctions and $\mathcal{E}$ represents the set of directed road segments encoded in a SUMO-compatible XML network file.

Algorithm 2: Edge identifier extraction.

Require: Network XML file N_{xml}
Ensure: Set of valid edges $\mathcal{E}_{\text{valid}}$
Parse N_{xml} into an XML tree structure
for all edge elements $e \in \mathcal{E}$ **do**
  Retrieve identifier $id(e)$
  if $id(e)$ satisfies directionality and semantic constraints **then**
    $\mathcal{E}_{\text{valid}} \leftarrow \mathcal{E}_{\text{valid}} \cup \{id(e)\}$
  end if
end for
return $\mathcal{E}_{\text{valid}}$

5.2.2 *Edge identifier extraction*

The purpose of this step is to find the edges that can be crossed for random origin-destination sampling. This abstraction makes it possible to consistently sample network elements, no matter how big or complicated the topology is, and it is done following Algorithm 2.

5.2.3 *Removal of pre-existing mobility entities*

To ensure statistical independence across simulation runs, all previously defined mobility entities are eliminated from the route configuration file before demand generation.

$$\mathcal{T} = \{\text{vehicle}, \text{pedestrian}\}, \tag{2}$$

Where $\mathcal{T}$ denote the entity types. The work done in this section is explained in details in Algorithm 3

Algorithm 3: Entity purging procedure.

Require: Route file R, entity type $\tau \in \mathcal{T}$
Ensure: Cleaned route file R^*
Load all lines from R
for all lines ℓ in R **do**
  if ℓ corresponds to entity type τ **then**
    Discard ℓ
  else
    Retain ℓ
  end if
end for
Rewrite retained lines to R^*

5.2.4 *Stochastic trip and pedestrian demand generation*

Let N denote the total number of mobility entities to be generated. Two admissible edge subsets are defined as

$$\mathcal{E}_{\text{start}}, \mathcal{E}_{\text{end}} \subseteq \mathcal{E}_{\text{valid}}, \tag{3}$$

corresponding to feasible origins and destinations.

For each entity $i \in \{1, \ldots, N\}$, an origin–destination pair is sampled such that

$$e_{\text{start}}^{(i)} \neq e_{\text{end}}^{(i)}. \tag{4}$$

5.2.5 *Entity instantiation model*

Each mobility entity is instantiated as follows:

Vehicle Model

$$\text{trip}_i = \langle id_i, t_i^{\text{dep}}, e_{\text{start}}^{(i)}, e_{\text{end}}^{(i)} \rangle. \tag{5}$$

Pedestrian Model

$$\text{person}_i = \langle id_i, t_i^{\text{dep}}, \text{walk}(e_{\text{start}}^{(i)}, e_{\text{end}}^{(i)}) \rangle. \tag{6}$$

In the mobility entity instantiation models, the index $i \in \{1, 2, \ldots, N\}$ is the ID of a single mobility entity, and N is the total number of entities created in the simulation scenario. For vehicle-based mobility, trip_i refers to the i-th vehicle trip, and for walking-based mobility, person_i refers to the pedestrian entity that goes with it.The id_i symbol stands for a unique number that is given to the i-th mobility entity. This identifier makes sure that entities can be clearly referenced and tracked during the simulation process.The parameter t_i^{dep} tells you when the i-th entity will leave, in simulation seconds, and when it will become active in the traffic network.

The letters $e_{\text{start}}^{(i)}$ and $e_{\text{end}}^{(i)}$ stand for the starting and ending edges of the i-th mobility entity. Both edges are chosen from predefined sets of acceptable edges, with $\mathcal{E}_{\text{start}}$ for origins and $\mathcal{E}_{\text{end}}$ for destinations. These sets are based on the network topology. To stop degenerate trips and make sure that mobility trajectories are meaningful, the condition $e_{\text{start}}^{(i)} \neq e_{\text{end}}^{(i)}$ is enforced.

The ordered tuple $\langle \cdot \rangle$ represents the combination of the various structured features of each mobility entity (such as identifying characteristics, time characteristics and location characteristics). When a mobility entity is created, it is incorporated into the SUMO simulation route file; this process activates the mobility entity in the SUMO execution environment and allows it to engage in subsequent traffic dynamics.

5.2.6 *Closed-Loop Simulation Execution*

The simulation evolves in discrete time steps $k \in \mathbb{N}$ and follows the steps explained in Algorithm 4.

Algorithm 4: Simulation control loop.

```
Initialise connection
for k = 1 to k_max do
    Advance simulation by one step
    Query active vehicle count V(k)
    Query active pedestrian count P(k)
    if V(k) = 0 and P(k) = 0 and k > 1 then
        break
    end if
end for
Close connection
```

5.2.7 *Reproducibility and Scalability Considerations*

The proposed system meets strict methodological standards due to its provision of a generic abstract method that allows replication and portability to multiple simulation environments. A statistically valid demand-generating model ensures traffic patterns generated with this framework are representative of real-world use cases and are also free from bias in the selection of origin and destination locations, both of which can impact traffic volume. The integration of near real-time simulation management capabilities allows the fine monitoring of simulations and affords the option for immediate correction during the execution of the simulation. In addition, since the architecture is highly scalable, it can be easily applied to urban transportation networks of varying sizes and complexity levels. Collectively, these features support the use of DRL-PVTSM system for developing additional extensions that involve the integration of optimisation-based control methods, advanced algorithms, and multimodal traffic analyses.

6 Results

Analysis of the experimental results reveals that increasing network architectural complexity has multi-dimensional effects on traffic performance. While average vehicle waiting time decreased from 54.35 seconds in the 1×4 network to 44.35 seconds in the 4×4 network—representing an 18.4% reduction—average trip duration exhibited a slight increase from 163.10 seconds to 171.21 seconds, reflecting longer travel distances in more complex networks. This trade-off suggests that the DRL-PVTSM framework effectively prioritizes the reduction of stop delays, a key factor in user experience, even as total path length increases. Car delays at intersections exhibited a similar trend, tending to decrease with increased network complexity. The most favourable results of our research on time loss were from the 3 × 3 configuration of road networks as it relates to pedestrian movement.

Additionally, our findings also indicate a positive relationship exists between 1) the level of complexity in the architecture of the road network, 2) the level of speed that cars moved at, and 3) the duration of travel of cars through the network. The addition of more intersections to the road network resulted in cars travelling at higher speeds and taking longer to get from one point to another. It looks like creating more intersections means more vehicle networks flow dynamics. To be precise and clear in this point we can say that there are so many different ways that design can change the way a network works.When you build a new network, you will affect the amount of time it takes for a vehicle to get to the destination and the enjoyment/comfort level of a person traveling along that network.

7 Future directions

The future of technology will be marked by advances in 6G wireless communications capabilities, as well as the growth of EC, which will significantly enhance our ability to collect location-based data on where pedestrians are located and to track them over time (i.e., how they are moving) as they traverse through urban environments [18]. The use of innovative and state-of-the-art methods will help us solve the existing problems with the usability and performance of many urban transit and transportation systems.

There is an opportunity to further examine how UAVs combined with AI may be utilised in future research on intelligent drone networks to improve monitoring of traffic conditions. It would be possible to develop an optimal and real-time monitoring regime that not only tracks vehicular movement but also provides for an opportunity to improve street-level efficiency and, thereby, enhance urban mobility based on the functionality of both modern UAV platforms and intelligent algorithm functionality provided through AI.

The concept of smart crosswalks that integrate sensor technology is another fascinating area of future study. Smart crosswalks equipped with sensor technology would allow for the generation and processing of data related to both pedestrian and vehicular traffic and would provide opportunities for the intelligent analysis of that data, allowing for automatic adjustment of traffic signal synchronisation

to create safer and more efficient pedestrian crossings and for the better overall management of urban traffic.

Multi-modal integration is an exciting future research opportunity and involves utilising AI algorithms to provide a comprehensive urban mobility solution [20]. For example, AI algorithms can create a fully integrated urban mobility system that integrates the use of all modes of transportation (cars, bicycles, walkers, buses, etc.) into one seamless transportation network. By intelligently developing and deploying comprehensive AI algorithms to account for all aspects of the environment in which many modes of transportation operate together, we can create a more comprehensive, efficient and effective transportation system.

There are many possibilities for diverse types of research and creativity within intelligent urban mobility, and the opportunities presented by new technologies, including 6G, EC, UAVs, and AI, will support the enhancement of pedestrian safety, the improvement of the control of traffic flow, and the development of a more interconnected and sustainable urban transportation system. This is just one aspect of potential work to be accomplished and demonstrates why an interdisciplinary approach is vital in considering urban mobility and transportation's many complex challenges moving forward.

8 Conclusion

In this work, we developed a decentralised traffic management system that resolves a basic disadvantage of DRL-based approaches which is that they are primarily vehicle-centric of transportation. Using the anticipated capabilities that will be available through the 6G network, we created a DRL-PVTSM model that allowed pedestrian dynamics to be integrated into how variable signal control is executed in real-time. This proposed model allows a balanced optimisation between preserving traffic efficiency and ensuring pedestrian safety, while preserving both scaling and computation efficiency at the same time. In our proof-of-concept simulation of both structured and random urban networks, we demonstrated statistically significant improvements when compared to fixed-time, vehicle-only adaptive, and centralised-control approaches. Reduced waiting times for pedestrians, reduced delays or vehicle downtime and lower overall system congestion all indicate that by using a pressure-based reward mechanism and a decentralised learning method, we were able to achieve our goal of effectively supporting the needs of both pedestrians and vehicles with our new approach. The data also indicates that increasing the number of vehicles in a distributed network does not degrade performance, which confirms and supports our conclusion that the proposed approach is appropriate to be used in large-scale urban environments. In addition to the immediate performance improvements provided, we have developed a model that allows us to treat pedestrians equally and not simply as a limiting condition, thereby creating the basis for developing intelligent urban transportation systems utilising these new technologies. With the development of wireless networks, edge intelligence, and next-generation sensor technology, our proposed DRL-PVTSM will allow urban environments to utilise the most efficient and equitable ways of managing mixed vehicle traffic.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Bao, Z., Ng, S. T., Yu, G., Zhang, X., & Ou, Y. (2023). The effect of the built environment on spatial-temporal pattern of traffic congestion in a satellite city in emerging economies. *Developments in the Built Environment, 14*, 100173. [CrossRef]

[2] U.S. House Committee on Transportation and Infrastructure, Democrats. (2011, September 27). *TTI Report: Cost of Congestion More Than $100 Billion*. Retrieved from https://democrats-transportation.house.gov/news/press-releases/tti-report-cost-of-congestion-more-than-100-billion

[3] Agarwal, I., Singh, A., Agarwal, A., Mishra, S., Satapathy, S. K., Cho, S. B., ... & Mohanty, S. N. (2024). Enhancing road safety and cybersecurity in traffic management systems: Leveraging the potential of

reinforcement learning. *IEEE Access, 12*, 9963-9975. [CrossRef]

[4] Nam, G. S., Yang, Q., & Yoo, S. J. (2025). Joint Optimization of Vehicle and Pedestrian Traffic Signals Using Multi-Objective Deep Reinforcement Learning. *IEEE Transactions on Intelligent Transportation Systems, 27*(1), 501-520. [CrossRef]

[5] Telikani, A., Sarkar, A., Du, B., & Shen, J. (2024). Machine Learning for UAV-Aided ITS: A Review With Comparative Study. *IEEE Transactions on Intelligent Transportation Systems, 25*(11), 15388-15406. [CrossRef]

[6] Anis Oukebdane, M., Shahen Shah, A. F. M., Baharul Islam, M., Ekoru, J., & Madahana, M. (2025). Hybrid Model for 6G Network Traffic Prediction and Wireless Resource Optimization. *IEEE Access, 13*, 142129-142139. [CrossRef]

[7] Anis Oukebdane, M., Shahen Shah, A. F. M., Kalam Azad, A., Ekoru, J., & Madahana, M. (2025). Unraveling the Nexus of ML and 6G: Challenges, Opportunities, and Future Directions. *IEEE Access, 13*, 114934-114958. [CrossRef]

[8] Guo, Q., Li, L., & Ban, X. J. (2019). Urban traffic signal control with connected and automated vehicles: A survey. *Transportation research part C: emerging technologies, 101*, 313-334. [CrossRef]

[9] Hassan, M. A., Elhadef, M., & Khan, M. U. G. (2023). Collaborative Traffic Signal Automation Using Deep Q-Learning. *IEEE Access, 11*, 136015-136032. [CrossRef]

[10] Chen, T., Zhang, K., Giannakis, G. B., & Başar, T. (2022). Communication-Efficient Policy Gradient Methods for Distributed Reinforcement Learning. *IEEE Transactions on Control of Network Systems, 9*(2), 917-929. [CrossRef]

[11] Qadri, S. S. U. H. S. M., Gökçe, M. A., & Öner, E. (2020). State-of-art review of traffic signal control methods: challenges and opportunities. *European Transport Research Review, 12*(1), 55. [CrossRef]

[12] Gregurić, M., Vujić, M., Alexopoulos, C., & Miletić, M. (2020). Application of deep reinforcement learning in traffic signal control: An overview and impact of open traffic data. *Applied Sciences, 10*(11), 4011. [CrossRef]

[13] Koch, L., Brinkmann, T., Wegener, M., Badalian, K., & Andert, J. (2023). Adaptive Traffic Light Control With Deep Reinforcement Learning: An Evaluation of Traffic Flow and Energy Consumption. *IEEE Transactions on Intelligent Transportation Systems, 24*(12), 15066-15076. [CrossRef]

[14] Haydari, A., & Yılmaz, Y. (2022). Deep Reinforcement Learning for Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems, 23*(1), 11-32. [CrossRef]

[15] Kolat, M., Kővári, B., Bécsi, T., & Aradi, S. (2023). Multi-agent reinforcement learning for traffic signal control: A cooperative approach. *Sustainability, 15*(4), 3479. [CrossRef]

[16] Xu, K., Huang, J., Kong, L., Yu, J., & Chen, G. (2022). PV-TSC: Learning to control traffic signals for pedestrian and vehicle traffic in 6G era. *IEEE Transactions on Intelligent Transportation Systems, 24*(7), 7552-7563. [CrossRef]

[17] Li, Y., Zhao, J., Zhang, G., & Shen, J. (2026). DRL-Based Robust Adaptive Traffic Signal Control with Low V2I Penetration Rate under Mixed Environment. *IEEE Transactions on Vehicular Technology*. [CrossRef]

[18] Chen, X., Wang, X., Zhao, W., Wang, C., Cheng, S., & Luan, Z. (2025). Hierarchical deep reinforcement learning based multi-agent game control for energy consumption and traffic efficiency improving of autonomous vehicles. *Energy, 323*, 135669. [CrossRef]

[19] Pan, T. (2023). Traffic light control with reinforcement learning. *arXiv preprint arXiv:2308.14295*.

[20] Bernárdez, G., Suárez-Varela, J., López, A., Shi, X., Xiao, S., Cheng, X., ... & Cabellos-Aparicio, A. (2023). Magnneto: A graph neural network-based multi-agent system for traffic engineering. *IEEE Transactions on Cognitive Communications and Networking, 9*(2), 494-506. [CrossRef]

[21] Sunil, R., Mer, P., Parmar, P., & Khan, N. (2025). Fusion of Emerging Technologies for 6G-Enabled Secure Smart City. *Security Paradigms in 6G Smart Cities and IoT Ecosystems*, 229-251.

[22] Eclipse Foundation. (n.d.). Eclipse SUMO: Simulation of Urban Mobility. Retrieved December 27, 2025, from https://www.eclipse.dev/sumo/

[23] German Aerospace Center (DLR). (n.d.). *SUMO User Documentation*. Retrieved October 15, 2025, from https://sumo.dlr.de/docs/

[24] Vinitsky, E., Kreidieh, A., Le Flem, L., Kheterpal, N., Jang, K., Wu, C., ... & Bayen, A. M. (2018, October). Benchmarks for reinforcement learning in mixed-autonomy traffic. In *Conference on robot learning* (pp. 399-409). PMLR.

[25] Abdulhai, B., Pringle, R., & Karakoulas, G. J. (2003). Reinforcement Learning for Adaptive Traffic Signal Control. *Journal of Transportation Engineering, 129*(3), 278-285. [CrossRef]

[26] Tataria, H., Shafi, M., Molisch, A. F., Dohler, M., Sjöland, H., & Tufvesson, F. (2021). 6G wireless systems: Vision, requirements, challenges, insights, and opportunities. *Proceedings of the IEEE, 109*(7), 1166-1199. [CrossRef]

[27] Liang, X., Du, X., Wang, G., & Han, Z. (2019). A deep q learning network for traffic lights' cycle control in vehicular networks. *IEEE Transactions on Vehicular Technology, 68*(2), 1243-1253. [CrossRef]

[28] Shirazi, M. S., & Morris, B. T. (2016). Vision-Based Vehicle and Pedestrian Tracking of Intersection Videos. *International Journal on Artificial Intelligence Tools, 25*(05), 1640004. [CrossRef]

[29] Panahi, F. H., & Panahi, F. H. (2024). Unmanned aerial vehicles toward intelligent transportation systems. *Interconnected Modern Multi-Energy Networks and Intelligent Transportation Systems: Towards a Green Economy and Sustainable Development*, 379-399. [CrossRef]

[30] Oukebdane, M. A., & Shahen Shah, A. F. M. (2025). Computer Vision-Powered 6G Networks: Technologies, Applications, and Challenges. *ICCK Transactions on Mobile and Wireless Intelligence, 1*(1), 19-31. [CrossRef]

[31] Akyildiz, I. F., Kak, A., & Nie, S. (2020). 6G and Beyond: The Future of Wireless Communications Systems. *IEEE Access, 8*, 133995-134030. [CrossRef]

[32] Rappaport, T. S., Xing, Y., Kanhere, O., Ju, S., Madanayake, A., Mandal, S., ... & Trichopoulos, G. C. (2019). Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond. *IEEE Access, 7*, 78729-78757. [CrossRef]

[33] Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y. P., Hilbrich, R., ... & Wießner, E. (2018, November). Microscopic traffic simulation using sumo. In *2018 21st international conference on intelligent transportation systems* (*ITSC*) (pp. 2575-2582). IEEE. [CrossRef]

[34] Behrisch, M., Bieker, L., Erdmann, J., & Krajzewicz, D. (2011). SUMO–simulation of urban mobility: an overview. In *Proceedings of SIMUL 2011, the third international conference on advances in system simulation*. ThinkMind.

[35] Zhang, G., Huang, H., & Chang, F. (2025). TS-PVL: Two-Stage Deep Reinforcement Learning-Based Traffic Light With Pedestrian-Vehicle Control in Mixed-Autonomy Traffic. IEEE Internet of Things Journal, 12(15), 31001-31014. [CrossRef]

[36] Wang, K., Shen, Z., Lei, Z., Liu, X., & Zhang, T. (2025). Toward Multi-Agent Reinforcement Learning Based Traffic Signal Control Through Spatio-Temporal Hypergraphs. *IEEE Transactions on Mobile Computing, 24*(9), 8258–8271. [CrossRef]

[37] Jiang, S., Huang, Y., Jafari, M., & Jalayer, M. (2022). A Distributed Multi-Agent Reinforcement Learning With Graph Decomposition Approach for Large-Scale Adaptive Traffic Signal Control. *IEEE Transactions on Intelligent Transportation Systems, 23*(9), 14689-14701. [CrossRef]



Mohammed Anis Oukebdane received the B.Sc. degree in telecommunication, Department of electrotechnical engineering and M.Sc. degree in Networks and telecommunications engineering from the University of Mustapha Stambouli, Algeria, in 2018 and 2020, respectively. He is currently pursuing Ph.D. degree in electronics and communication engineering in Yildiz Technical University, Turkey. He is working in AI and Next-generation Wireless Communication Laboratory (ANWCL). His current research interests include wireless communications, RIS, 6G, FANETs, UAVs' automation systems, and cross-layer design.