

Review Article

DRL-Based Intelligent Resource Allocation for Diverse QoS in 5G and toward 6G Vehicular Networks: A Comprehensive Survey

Hoa TT. Nguyen ¹, **Minh T. Nguyen** ², **Hai T. Do**,² **Hoang T. Hua**,² and **Cuong V. Nguyen**³

¹*Sejong University (SJU), Republic of Korea*

²*Thai Nguyen University of Technology (TNUT), Vietnam*

³*Thai Nguyen University of Information and Communication Technology (ICTU), Vietnam*

Correspondence should be addressed to Minh T. Nguyen; tuanminh.nguyen@okstate.edu

Received 9 May 2021; Accepted 2 August 2021; Published 17 August 2021

Academic Editor: Stefan Panic

Copyright © 2021 Hoa TT. Nguyen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The vehicular network is taking great attention from both academia and industry to enable the intelligent transportation system (ITS), autonomous driving, and smart cities. The system provides extremely dynamic features due to the fast mobile characteristics. While the number of different applications in the vehicular network is growing fast, the quality of service (QoS) in the 5G vehicular network becomes diverse. One of the most stringent requirements in the vehicular network is a safety-critical real-time system. To guarantee low-latency and other diverse QoS requirements, wireless network resources should be effectively utilized and allocated among vehicles, such as computation power in cloud, fog, and edge servers; spectrum at roadside units (RSUs); and base stations (BSs). Historically, optimization problems have mostly been investigated to formulate resource allocation and are solved by mathematical computation methods. However, the optimization problems are usually nonconvex and hard to be solved. Recently, machine learning (ML) is a powerful technique to cope with the complexity in computation and has capability to cope with big data and data analysis in the heterogeneous vehicular network. In this paper, an overview of resource allocation in the 5G vehicular network is represented with the support of traditional optimization and advanced ML approaches, especially a deep reinforcement learning (DRL) method. In addition, a federated deep reinforcement learning- (FDRL-) based vehicular communication is proposed. The challenges, open issues, and future research directions for 5G and toward 6G vehicular networks, are discussed. A multiaccess edge computing assisted by network slicing and a distributed federated learning (FL) technique is analyzed. A FDRL-based UAV-assisted vehicular communication is discussed to point out the future research directions for the networks.

1. Introduction

The 5G new radio (NR) is driven by the demand for the large volume of data due to the burst growth of cellular mobile devices and vehicles [1]. It allows numerous different applications to make lives easier, smoother, and more comfortable. In the intelligent transportation system (ITS), autonomous driving, and smart city areas, 5G is expected to improve safety, increase experience comfort, reduce traffic congestion, and lower air pollution [2]. Vehicles will be highly connected with the aid of the ubiquitous wireless network at anytime and anywhere [3]. Thus, the 5G vehicular

network becomes heterogeneous due to different applications. Accordingly, the quality of services (QoS) is diverse. One critical requirement in the vehicular network to guarantee the safe transportation is an extremely low-latency communication. To satisfy these QoS demands, resource management approaches such as computation power at cloud, fog, and edge servers; spectrum allocation at roadside units (RSUs); and base stations (BSs) for vehicular user equipment (VUE) have been extensively investigated. Normally, optimization problem approaches are formulated for resource management such as computation power minimization, sum-rate maximization, and latency minimization. In a

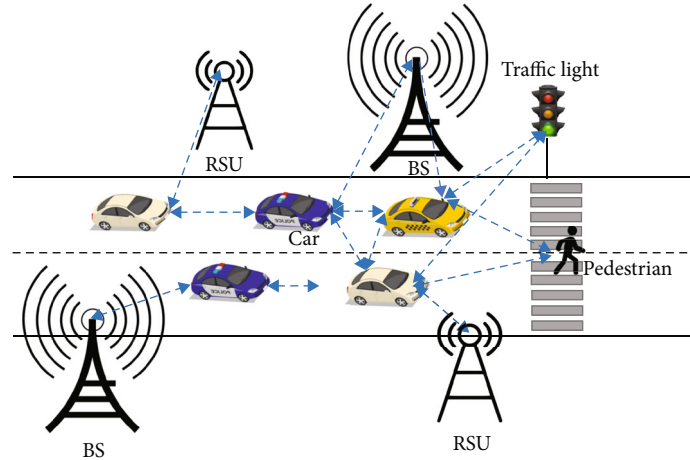


FIGURE 1: A 5G heterogeneous vehicular network.

few simple cases, a convex optimization is enough to achieve these objectives. However, in reality, most of the formulated optimization problems for wireless resource management are strongly nonconvex and nondeterministic polynomial-time hardness (NP-hard). Due to the hard mathematical computation, there are no effective and strong enough algorithms to find the optimal or even suboptimal points. In addition, the vehicular network comes with a lot of new service applications and the natural-mobility characteristics. Accordingly, the data becomes extremely big and hard to analyze. Thus, a new strong computation method needs to be deployed to cope with these challenges. A machine learning (ML), especially deep reinforcement learning (DRL), is an emerging algorithm to work with big data and data analysis, which can effectively support in resource management for vehicular networks. More than that, a distributed learning approach such as federated deep reinforcement learning (FDRL) allows DRL algorithm learning without sharing vehicles' dataset. Thus, the latency is reduced and the privacy is guaranteed. In order to provide connectivity everywhere and every time, UAVs are deployed in vehicular networks and act as flying BSs. Therefore, a FDRL-based UAV-assisted 5G and toward 6G vehicular networks are proposed.

5G and 6G vehicular networks become heterogeneous, including vehicle-to-vehicle (V2V) links, vehicle-to-infrastructure (V2I) links, vehicle-to-pedestrian (V2P) links, and vehicle-to-everything (V2X) links [4]. More specifically, a 5G heterogeneous vehicular network is illustrated in Figure 1. A dedicated short range communication (DSRC) is designed as a two-way short-range communication channel among vehicles. A cellular-based vehicular network comprises macro-BSs and RSUs to provide a wider communication range and high data rate services for VUEs. More specifically, in the 5G vehicular network, vehicles are connected to each other, to the infrastructure, and to pedestrians on the ground to support intelligent transportation purposes. Unlike 5G, the 6G AI-enabled vehicular network does not only keep the same communication architecture of the intelligent transportation but also allows communications among space-air-ground (SAG), even underwater vehicles, which is illustrated in Figure 2. In space, there are

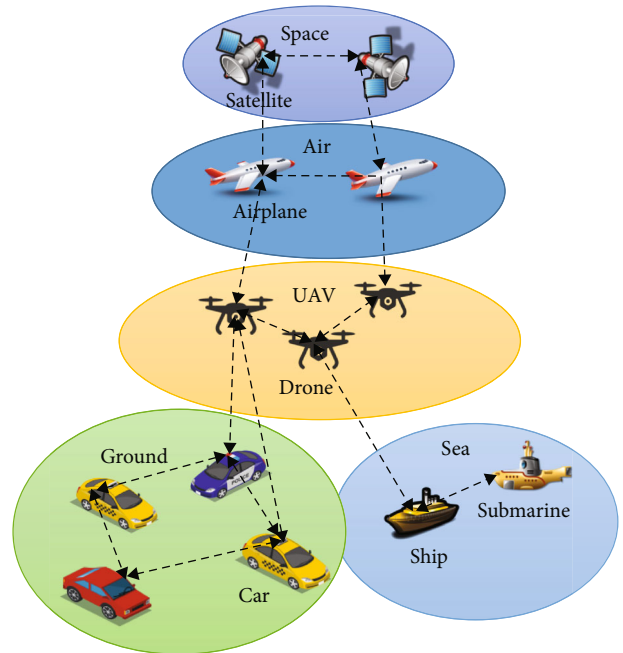


FIGURE 2: A 6G heterogeneous vehicular network.

satellite communications and communications between the satellite level and the air level (e.g., airplanes). These airplanes are connected to each other to transmit data or information. Below the air level is UAV communication. These UAVs have two missions, which includes communication to the airplanes and communication to the ground level (e.g., BSs, RSUs, and autonomous driving cars). At the ground level, there are V2V, V2I, and V2X communications to ensure the safe transportation. Moreover, at the sea level, unmanned underwater vehicles (e.g., submarine) and ships are connected to each other and connected to UAVs in case of disaster emergency rescue. According to the aforementioned heterogeneous 5G and 6G vehicular networks, their QoS requirements become diverse.

5G NR supports new diverse QoS demands targeting ultra-reliable low-latency communication (URLLC),

enhanced mobile broadband (eMBB), and massive machine-type communications (mMTC) [5]. The URLLC service requires 99.999% reliable transmission within 1 ms end-to-end (E2E) latency with larger support of up to 1 million devices/km². After completion in investigation of 5G NR, the 6G network will be investigated for the future evolution of a network intelligentization [6] and is expected to be deployed by 2030 [7]. Some predicted scenarios for the 6G network are described as an improvement of the 5G network. A further-enhanced mobile broadband (FemBB) with a peak throughput of 1 Tb/s (1000 times larger than the 5G network) [8] is necessary for the 6G network. The delay requirement defined as event-defined ultra-reliable and low-latency communication (EDURLLC) is reduced by at least 10 times (from 1 ms to less than 50 ns) [9]. The reliability needs to be guaranteed at 99.99999% [10] to support unmanned systems such as autonomous driving and unmanned aerial vehicles (UAVs). Moreover, spectrum and energy efficiency must be achieved over 10 times compared to the 5G network. The 6G network requires connection density of up to 1000 times to reach 10⁷–10⁸ km² [11], and mobility requirement is enhanced from 500 km/h to subsonic 1000 km/h (airplane) [12]. Further, it demands extremely low-power communication (ELPC) and long-distance and high mobility communication (LDHMC) [13]. Therefore, new application scenarios towards 2030 in the 6G network are intelligent life, intelligent production, and intelligent society [14].

To satisfy the diverse QoS requirements in the 5G vehicular network, computation power needs to be effectively utilized at cloud, fog, and edge servers and spectrum resource should be fairly allocated to VUEs at RSUs and BSs [15]. Thus, an optimization problem in a joint of computation power and radio resource utilization needs to be addressed to support diverse QoS requirements. Most of optimization problems are formulated to maximize sum-rate and energy efficiency and minimize E2E latency under the QoS constraints. However, these optimization problems are nonconvex and hard to solve due to mixed-integer nonlinear programming (MINLP). In addition, with the burst increase in the number of VUEs, the problem of big data and data analysis becomes complex. Thus, ML is proposed to cope with these challenges. According to the learning ways to build the models, ML can be separated into four different categories [16]. Firstly, a supervised learning is a type of algorithm that works with labeled data to find a good function of mapping the inputs and the outputs. Secondly, an unsupervised learning, such as deep learning (DL), is a technique that works with unlabeled data to find the common traffic patterns and then divides them into clusters. Thirdly, a semi-supervised learning is between supervised and unsupervised learning algorithms and can be divided into transductive and inductive learning. Lastly, a reinforcement learning (RL) is a technique to find a policy to maximize the sum of rewards. Recently, there are two ML approaches in solving resource management problems. The first approach is a DL-assisted optimization problem. The objective function in the optimization problem is treated as a loss function in supervised learning approach. Then, stochastic gradient descent is employed to find the suboptimal and optimal

points. However, in the heterogeneous vehicular network, how to obtain the training dataset for training an ML algorithm is a big problem. Thus, the other approach is deep reinforcement learning- (DRL-) based resource allocation. DRL supports decision making in resource allocation at RSUs and BSs by designing a reward signal related to the ultimate goal. Then, the learning algorithm automatically finds a gradient decent solution to manage resource in the vehicular network.

A federated learning (FL) method can deal with the lack of privacy due to all of the raw data in vehicles transmitted to the central server for computing and processing. It is a distributed machine learning method and first proposed by Google [17]. It allows multiple parties (e.g., VUEs) to jointly train a model at a BS or a RSU while mitigating the privacy risk and reducing the latency. In traditional ML approaches, ML algorithms are employed at the central servers. Then, all the raw data from each vehicles are transmitted to the central servers for computing and making decisions to allocate radio and power resources for individual vehicles. Even though, via this approach, the central servers have the overview of all their vehicles, privacy issues are not guaranteed. In the contrast, for a FL approach, the same ML algorithms are employed at both the central servers and multiple vehicles. At first, global ML models at the center servers initialize global parameters. Then, they choose vehicular participants to incorporate for training the global models and send these parameters to them. The participants use the received parameters to train their local models based on their own local dataset. Instead of transmitting all the raw data to the center servers, after a predefined training period, vehicles only send the update parameters to the center servers. Then, the central servers aggregate all the new parameters via an algorithm such as federated averaging (FedAvg) and use the updates to train their own models. These steps are iterative until the global models are satisfied. Thus, the privacy is guaranteed and the latency is reduced.

In order to achieve the new stringent latency requirement, immediate vehicular connectivities need to be satisfied. However, due to the vehicular traffic dynamics, the traditional infrastructures (e.g., macro BSs, RSUs) can not immediately provide the connectivities. Thus, it leads to negative impacts on the vehicular connectivity performance such as weak communication links and broken links among V2V and V2I communication. To deal with these challenges, unmanned aerial vehicles (UAVs) are used as flying BSs to provide connectivities for vehicles. In the communication scenario, UAVs are equipped with wireless communication devices and other related electronics to serve the vehicular network. In the work of [18], UAVs act as flying BSs to deliver data from the vehicular sources to the vehicular destinations, or the far away BSs and RSUs. By this approach, the data delivery delay is reduced. More specifically, due to the advanced properties such as flexibility, mobility, and adaptive altitude, UAVs can provide wireless network connection to improve capacity, coverage, and energy efficiency to the vehicular network. However, due to the energy limitation of the battery in UAVs, there must be intelligent methods (e.g., transmit power and resource allocation strategy and

position selection) based on ML approaches to improve UAV capacity. Moreover, obtaining the dataset for the training phase in ML algorithms is hard due to the dynamics of the vehicular environment and UAVs. DRL is one of the best ML algorithms, which can learn without the dataset. More efficiently, to reduce latency and signaling overhead, a FL algorithm named federated deep reinforcement learning (FDRL) can be used.

The main problem addressed in this comprehensive survey is ML, especially DRL approach-based resource management techniques to guarantee the diverse QoS requirements in 5G and toward 6G vehicular networks. More specifically, the main contributions are to overview the recent advanced techniques such as mathematical computation, ML, DRL, and FDRL approaches for computation power in cloud, fog, and edge layers and resource allocation at RSUs, BSs, and UAVs in 5G and toward 6G vehicular networks summarized as follows:

- (i) Providing the diverse QoS requirements in 5G and 6G vehicular networks
- (ii) Reviewing the recent advanced techniques of traditional mathematical computation and ML and DRL approach-based resource management in cloud, fog, and edge servers at BSs, RSUs, and UAVs
- (iii) Analyzing the multiaccess edge computing with network slicing and distributed FL techniques, FDRL-based UAV-assisted communication to support the vehicular network
- (iv) Pointing out challenges, open issues and future research direction to guarantee the diverse QoS requirements in 5G and toward 6G vehicular networks

Along with the emerging research topic represented in this comprehensive survey, there are various related surveys as follows.

A survey of a self-organizing network in vehicular ad hoc networks (VANETs) is represented [19]. The issues of resource allocation related to the heterogeneous vehicular communication are discussed. Similarly, in the work in [20], a heterogeneous vehicular network (HetVNET) integrated with a cellular network in dedicated short-range communication is reviewed through a cross layer (i.e., medium access control (MAC) and network layer). However, a computation power and resource allocation in the physical layer based on ML and DRL algorithms are not investigated in these comprehensive surveys. Besides that, a visible light communication (VLC) usage is reviewed in vehicle applications [21]. The recent advanced techniques in VLC, their challenges, and solutions are analyzed. In [22], resource management is discussed in cloud computing for the vehicular network. The unique characteristics of vehicular nodes (i.e., high mobility, resource heterogeneity, and intermittent network connection) are included to cope with resource management. In [23], the authors represent a comprehensive survey on resource allocation for the two dominant vehicular

communication technologies including dedicated short-range communication (DSRC) and cellular vehicular networks (CelVNET). However, DRL-based resource management to improve the diverse QoS requirements are not included in these articles.

ML technique- (e.g., support vector machine (SVM), K-nearest neighbor (KNN), deep learning (DL), and deep Q-learning) based solutions for the vehicular network are discussed. The connected and autonomous vehicles (CAVs) and intelligent transportation system (ITS) supported by ML algorithms are reviewed [24]. In [25], an intelligent and secure toward 6G vehicular network with the support of ML algorithms is represented. However, the paper [24, 25] mainly discusses about security issues over the vehicular network while challenges of resource management (i.e., computation power and spectrum allocation) in both cloud, fog, edge computing, and RSUs; BSs are not mentioned in these comprehensive surveys. An AI approach-based resource management is represented [26]. An ML algorithm integrated with the cognitive radio vehicular ad-hoc network (CR-VANET) is reviewed. Even though the recent advancements in the amalgamation of prominent technologies and future research direction are given, resource management in both cloud, fog, and edge computing based on DRL and FDRL and assisted by UAVs to reduce latency for the vehicular network is not discussed in the survey.

Since the comprehensive overview of computation power in cloud, fog, and edge servers and spectrum allocation supported by ML, especially DRL, and FDRL algorithms, and assisted by UAVs in 5G and toward 6G vehicular networks has not received much attention. Therefore, in this comprehensive survey, a DRL-based computation power and resource allocation to guarantee the diverse QoS requirements is given.

The rest of this work is organized as follows. Section 2 provides the DRL background. The diverse QoS requirements in 5G and 6G vehicular networks are addressed in Section 3. Computational power and resource allocation approaches in cloud, fog, and edge computing for the 5G vehicular network are discussed in Section 4. In Section 5, UAV-assisted vehicular communication techniques are discussed with the support of the FDRL algorithm. Challenges, open issues, and future direction to manage resource are given in Section 6. Finally, conclusions are given in Section 7.

2. Background of Deep Reinforcement Learning

Deep reinforcement learning (DRL) is a machine learning algorithm. It consists of reinforcement learning (RL) and deep learning (DL). A sequential decision making is addressed via minimizing a reward while interacting with the unknown environment in every different application. Since the algorithm does not require many datasets for training itself, it is suitable with the dynamic environment characteristics in 5G and toward 6G vehicular networks.

2.1. Reinforcement Learning. Reinforcement learning (RL) is an adaptable algorithm which learns with the absence of a training dataset [27]. Thus, it automatically adapts to the

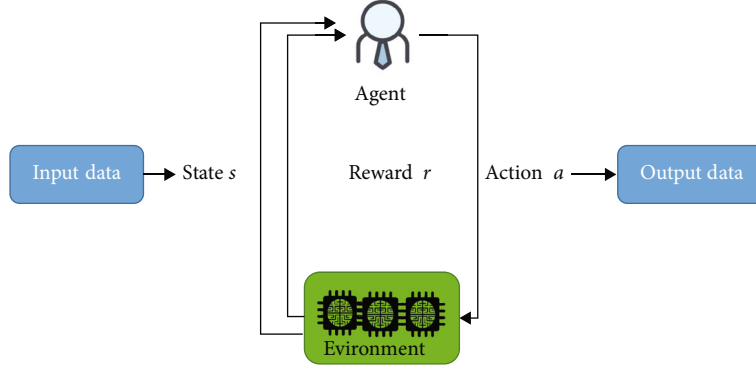


FIGURE 3: A reinforcement learning model.

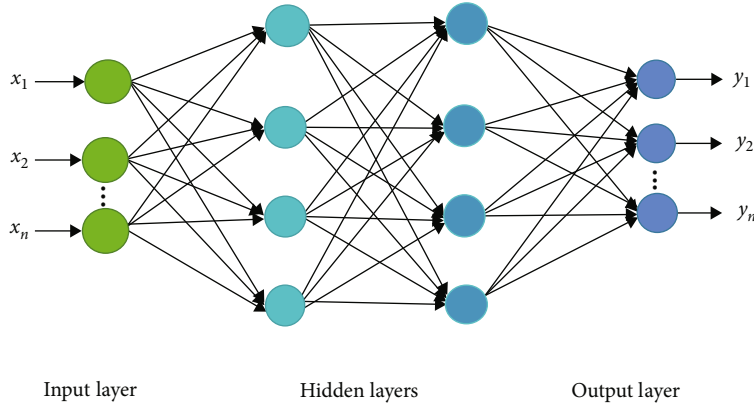


FIGURE 4: A fully connected deep network model.

new environment. Moreover, it allows online learning to support real-time signal communication between V2V, V2I, and V2X links [28]. Hence, it takes advantages in a dynamic vehicular network. RL has applications in healthcare [29], finance [30], and manufacturing [31]. In the autonomous driving area, the RL algorithm has been widely investigated in both academia and industry to support the stochastic vehicular environment due to time-varying service demands [32].

A basic RL model is illustrated in Figure 3. An agent is an entity, which performs an action a in an environment to gain a reward r . Hence, an environment in the physical world is in which the agent performs its action. A state s describes the current situation of the environment and all the possible actions that the agent can take. More specifically, at each time, the agent observes some representation of the environment state s by interacting with it. Then, the agent selects an action a from the action set \mathcal{A} . Following the action, the agent receives a reward r . One of the most common algorithms used in the field of RL is a Q-learning algorithm. The reward value in this algorithm is called Q value given by [33] as follows:

$$Q^*(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_{\{a'\}} Q(s', a') - Q(s, a)], \quad (1)$$

where $Q^*(s, a)$ and $Q(s, a)$ are the new Q value and the old Q value, respectively. α is the learning rate and $\gamma \in (0, 1]$ is the discount factor. The notation r denotes the reward. The term of $\max_{\{a'\}} Q(s, a')$ is the target.

2.2. Deep Learning. A deep learning (DL), known as a deep network, is based on artificial neural network (ANN). There are different existing DL algorithms such as the recurrent neural network (RNN), long short-term memory (LSTM), convolutional neural network (CNN), and deep neural network (DNN). Deep learning has many applications in autonomous driving [34], finance [35], and agriculture [36]. Since a DL model requires less formal statistical training, it can deal with the characteristics of the recent vehicular network. In addition, it has capability to detect nonlinear relationships between dependent and independent variables to improve the accuracy prediction.

A fully connected deep network model is illustrated in Figure 4. There are three types of layers, which comprises of an input layer, hidden layers, and output layer. The number of neuron cells in each layer and the number of hidden layers can be varied according to the different applications. One of the common algorithms used in the field of DL is an LSTM model. Its architecture is illustrated in Figure 5. In this model, an LSTM cell is considered as a neuron cell in the hidden layers of a fully connected deep network.

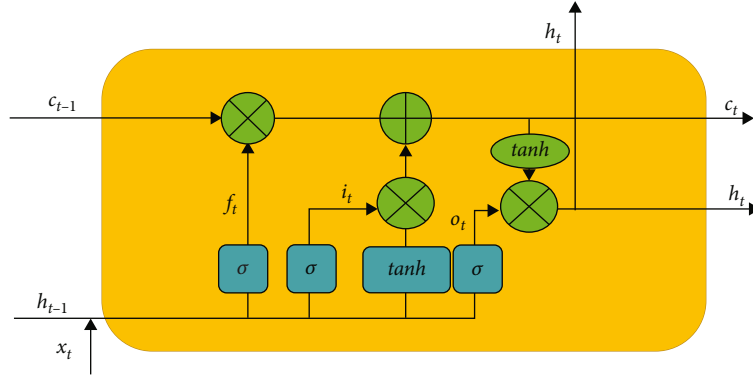


FIGURE 5: A long short-term memory model.

A specific example of handwriting recognition is described to explain Figure 5. There are several steps happening in a LSTM cell, given in [37]. In the first step, in order to predict the next word in a sentence based on the historical data, the data might be stored in the LSTM cell. In case of changing the predicted subject, the cell wants to forget the previous data. This process is performed in the forget gate given by at time t

$$f_t = \sigma(V_f x_t + U_f h_{t-1} + b_f), \quad (2)$$

where f and σ denote the forget gate's activation vector and the sigmoid activation function, respectively. V and U are the weight matrices. The notations x and h are the input vector and the hidden state vector. b denotes the bias vector. In the second step, the LSTM cell decides which data should be stored. This step must be done within two parts. The first part decides which data should be updated by sigmoid function σ at the input gate's activation vector i

$$i_t = \sigma(V_i x_t + U_i h_{t-1} + b_i). \quad (3)$$

The second part creates a vector of data \tilde{c} by using tanh activation function

$$\tilde{c}_t = \tanh(V_c x_t + U_c h_{t-1} + b_c). \quad (4)$$

The third step is to decide whether the cell should drop the data and the cell gets the new potential data c by

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t. \quad (5)$$

At the end, the LSTM cell decides what should be the output by performing two parts. The first part makes a decision of which data is going to the output by a sigmoid activation function

$$o_t = \sigma(V_o x_t + U_o h_{t-1} + b_o), \quad (6)$$

and the second part is to normalize the values of data between -1 and 1 given by

$$h_t = o_t \circ \sigma(c_t). \quad (7)$$

After prediction, the fully connected deep network with LSTM cells in hidden layers calculates the loss function (e.g., mean squared error) to validate the model.

2.3. Deep Reinforcement Learning. A deep reinforcement learning (DRL) is a computational approach of learning from an action. It can learn in the absence of a training dataset. It comprises of a DL algorithm and an RL algorithm to the agent learning the best actions in the virtual environment to attain its goal. DRL has many different existing applications (e.g., self-driving cars, industrial automation, trading and finance, and natural language processing).

Deep Q-learning is one of the most common algorithms used in the field of DRL. A deep Q-learning model is illustrated in Figure 6. The algorithm consists of a DL and a Q-learning algorithm. In the deep Q-learning, a DL network (e.g., RNN model) is used to approximate the Q value. The states $s \in \mathcal{S}$ are considered as the inputs of a DL network, and the Q values of all possible actions $a \in \mathcal{A}$ are generated as the outputs of the DL network.

Instead of updating Q values in every iteration as in the Q-learning algorithm, the objective of a deep Q-learning algorithm is to minimize a loss function L by updating parameter θ_k in every iteration k , given by [33] as follows:

$$L(s, a | \theta_k) = \left(r + \gamma \max_{\{a'\}} Q(s', a') - Q(s, a | \theta_k) \right)^2, \quad (8)$$

where θ_k is the weight and bias parameter needed to be updated in every iteration k .

3. QoS Requirements in the 5G and 6G Vehicular Networks

3.1. QoS Requirements in the 5G Vehicular Network. 5G is built to support three categories including ultra-reliable low-latency communication (URLLC), enhanced mobile broadband (eMBB), and massive machine-type communications (mMTC) [5]. The V2X application scenario was defined in the standard of the enhanced 5G V2X service. It supports advanced applications including extended sensor and state map sharing, vehicle platooning, remote driving, and advanced driving. 5G connectivity is built upon the

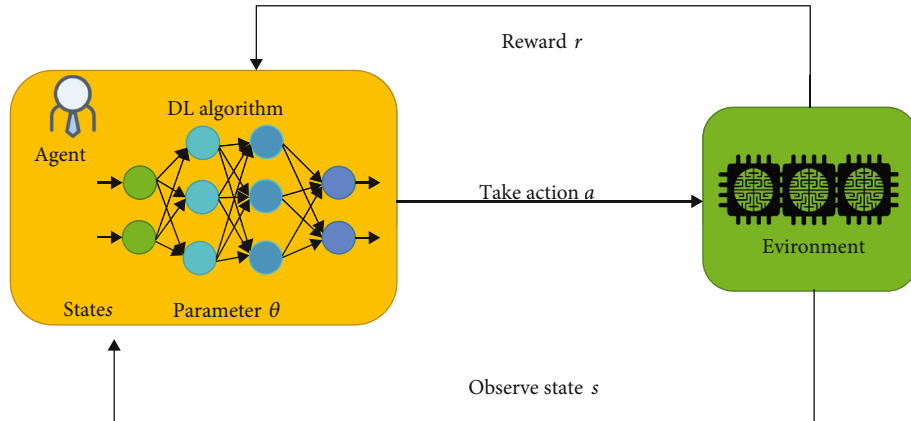


FIGURE 6: A deep Q-learning model.

packet data unit (PDU) session concept from information-centric networks. Each PDU session comprises of multiple QoS flows. It defines the finest granularity of QoS differentiation within a PDU session. In general, a QoS metric is defined as a set of parameters. It includes packet delay budget (PDB) (ms), packet error rate (PER) (%), and guaranteed and maximum bit (GMB) rate (kps). Vehicle platooning services have requirements of end-to-end (E2E) latency down to 10 (ms) with 99.99% reliability. On the other hand, advanced driving services [38] require communication channels, which must support the bitrate up to 1 (Gb/s) with the E2E latency down to 3 (ms) and reliability up to 99.999%.

3.2. QoS Requirements in the 6G Vehicular Network. 6G supports a fully connected and intelligent digital world [7]. It is expected to be deployed by 2030. The communication system is expected to associate with services, which have different requirements. One of these services is ultra-high speed with low-latency communications (uHSSLC). It reduces E2E latency to less than 1 (ms) [10] with more than 99.99999% reliability [11]. In the scenario of ultra-high data density (uHDD), 6G is expected to provide Gbps coverage everywhere. The coverage of the new environment such as the sky and sea is up to 10000 (km) and 20 (nautical miles) [39], respectively. Moreover, ubiquitous mobile ultra-broadband (uMUB) requires 1 (Tps) peak data rate. The 6G network will provide massive machine-type communication up to 10 (million devices/km²). In the other hand, autonomous driving has characteristics in ultra-high mobility (up to 1000 (km/h)) with terabytes generated per driving hour. In this scenario, 6G is expected to support reliability up to 99.99999% and E2E latency down to 1 (ms) [39].

4. Computation Power and Resource Allocation Techniques for the 5G Vehicular Network

According to the aforementioned 5G heterogeneous vehicular network, computation power and resource allocation have been widely investigated to guarantee the diverse QoS requirements. It includes stringent sensitive latency, high throughput, and massive connected vehicles. Among all of them, E2E latency is the most strict requirement in vehicular

communications. By effectively utilizing computation power and resource allocation, the latency can be guaranteed. Resource management approaches in vehicle cloud computing (VCC) [40], vehicle fog computing (VFC) [41], and vehicle edge computing (VEC) [42] are illustrated in Figure 7, which have been significantly studied to guarantee the desirable latency 1 ms. The recent advanced techniques of computation power and spectrum allocation are analyzed in cloud, fog, and edge computing via traditional optimization approaches, ML-based approaches, and DRL-assisted approaches. These techniques are represented in the following subsections.

4.1. Computation Power and Resource Allocation in Cloud Computing for the 5G Vehicular Network. A cloud computing is a network access model [43]. In other words, it is the delivery of computing services, including processing, computation power, and resources from vehicles to the cloud servers. Following the delivery, wireless network resources in vehicles are effectively saved. From that, the resources become available for sharing among other neighbor vehicles. A cloud computing system is illustrated in Figure 8. In this scenario, RSUs and macro-BSs act as gateways and are used to connect VUEs to their cloud servers while vehicles are moving. There are four layers in a cloud computing network for the vehicular network [44]. It consists of a perception layer, a coordination layer, an AI layer, and a smart application layer.

4.1.1. Optimization Theory Techniques. An approach of service-based resource block (RB) allocation to improve spectrum efficiency is proposed [45]. The services include basic safety services (beacon messages), advanced safety services (cameras), and traffic efficiency services (dynamic map update and intersection speed advisory). They are defined as user's service experiences. In this paper, the total bandwidth at the BSs is maximized via a formulated optimization problem under the QoS constraints. Similarly, in [46], spectrum and transmit power are effectively allocated to ensure the safety-critical information transmission for Internet of vehicle (IoV). The ergodic capacity of V2I links is maximized under latency constraints. Hence, the latency

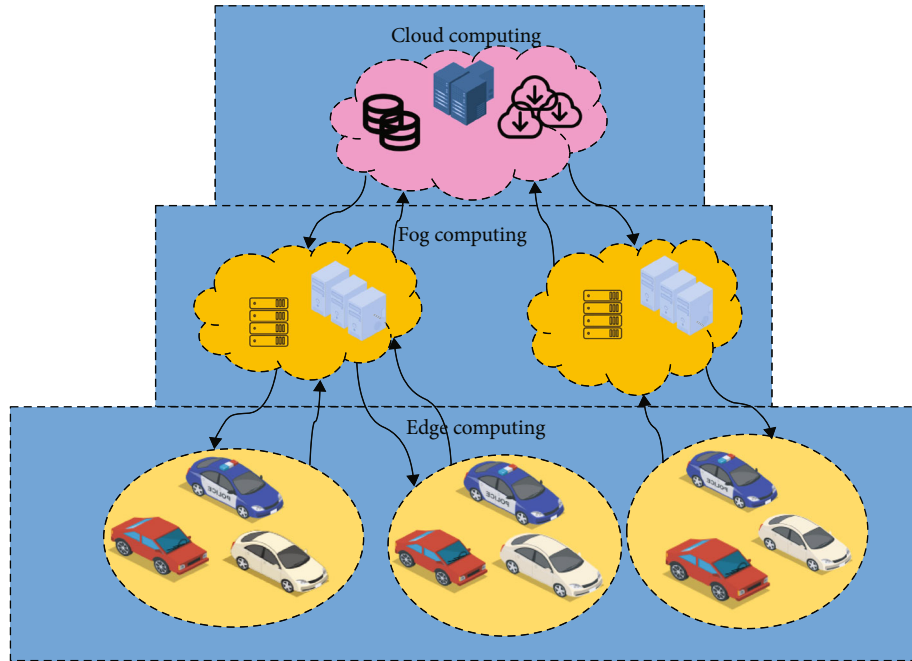


FIGURE 7: A computing model for the vehicular network.

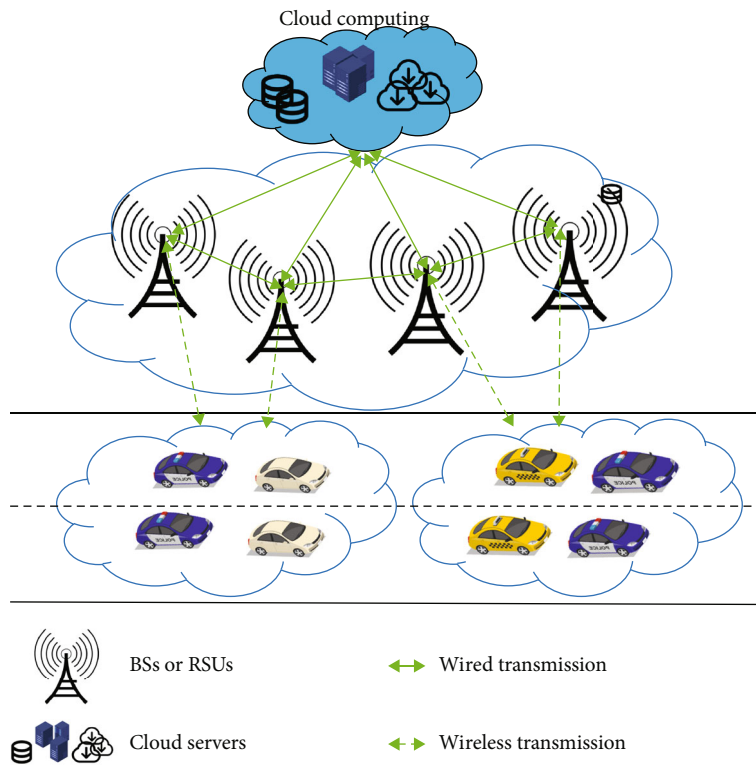


FIGURE 8: A cloud computing model for the vehicular network.

requirement of V2V links is achieved via the simulation results. In order to meet the ultra-reliable low-latency vehicular communication, a strategy of virtual cell member selection and time-frequency RB allocation is identified [47]. Specifically, a position-based user centric radio resource management is investigated. The results show that ultra-

reliable low-latency requirement is guaranteed in the vehicular network.

Besides E2E latency, reliability is guaranteed by effectively allocating the radio resource [48]. When vehicles appear in the out of coverage area, it leads to periodic and aperiodic V2V communication. In the article, a prediction

algorithm is used to predict the arrival rate of ad hoc services. Then, the radio resource defined as RBs is preallocated for the unexpected events inside the delimited out of coverage areas. In [49], an integration of a cloud computing and a vehicular network is studied. Vehicles share computation power, bandwidth, and storage resources to guarantee the URLLC requirement via an optimization problem. Then, a game theoretical approach is used to effectively allocate all these above resources. A good performance in the shared resources is achieved via the simulation results. Another approach, an optimal RB, an optimal power control, and optimal number of active antennas are considered to maximize energy efficiency in the vehicular network [50]. In addition, a short block length of channel codes is proposed to reduce transmission latency and to increase reliability. The simulation results show that the low-latency transmission and highly reliable communication are achieved.

Despite the good performance of low-latency high-reliable vehicular communication, complex computation of the optimization problems is a critical issue. Thus, ML algorithms are employed to tackle these challenges.

4.1.2. Supervised Learning-Assisted Techniques. A policy framework-based resource allocation is investigated to guarantee the diverse QoS requirements [51]. A policy framework, consisting of multiple network slices, is designed with the supports of software-defined networking (SDN) and network function virtualization (NFV). A network slice is an independent end-to-end (E2E) logical network. Then, multi-slices share a common physical infrastructure, including the radio access network (RAN), core network (CN), unmanned aerial vehicles (UAVs), and satellites. From that, the QoS requirements are guaranteed. Recently, a combination of a software-defined function with an ML algorithm is an emerging research topic in both mobile and vehicular networks. In [52], a network slice is defined through the SDN-enabled core network and SDN-enabled wireless data plane. In the article, a resource management module (RMM) is designed to estimate the radio resource for the individual slice. The concurrent neural network (CNN), deep neural network (DNN), and long short-term memory (LSTM) algorithms are used to classify the vehicular traffic patterns of vehicles. From the classification results, the same traffic VUEs are assigned according to their corresponding network slices. Then, the corresponding RSU or BS allocates radio resource slices to these network slices. The simulation results show that the latency and the other QoS requirements are guaranteed.

In some cases, network slices are defined as mobile virtual network operators (MVNOs) and share a common physical infrastructure of a mobile network operator (MNO). The transmit power and overhead of signaling in vehicles are reduced [53]. In this paper, an emerging distributed federated learning (FL) is studied to estimate the tail distribution of the queue length in the vehicular network. The queue length distribution is described by the extreme value theory (EVT). It is estimated by a maximum likelihood estimation (MLE) algorithm. Finally, the distributed FL is applied to reduce the signaling overhead. Accordingly, the optimal transmit power is used to release the queue. The simulation

shows that the ultra-reliable low-latency vehicular communication is guaranteed while the signaling overhead is reduced. An age of information- (AoI-) based resource allocation is investigated [54]. An AoI is defined as the time elapsed since the generation of the last received status update. This information is kept in VUEs for network analysis purposes. When these information are growing too large, it leads to information violation problem. In this paper, the tradeoff of reliability between maximizing the knowledge gains about the network and minimizing the probability that the AoI exceeds a predefined threshold is analyzed. After that, a Gaussian process regression (GPR) approach is studied to predict the future of AoI exceeding a predefined threshold.

Luckily, the supervised learning shows the good performance in the dynamic vehicular network. However, in the practical scenario, it is hard to obtain the training data for ML models in the dynamic environment. Hence, a DRL algorithm is proposed to train the model with the absence of training data.

4.1.3. DRL-Based Techniques. A DRL approach-based computational offloading algorithm for multilevel-vehicular cloud computing is investigated [55]. More specifically, an AoI-aware radio resource allocation is analyzed. Then, a discrete-time single-agent Markov decision process (MDP), one of the DRL algorithms, is applied to capture the AoI. From the observation, another single-agent MDP is applied to help the RSU allocating frequency bands for all VUEs. The results show that the total shared radio resources among vehicles are minimized while the low-latency high-reliable vehicular communication is guaranteed. Besides that, in order to have a safe traffic driving and to keep the stability among vehicles in a platooning problem for the vehicular network, a joining of the optimal radio resource and control problem is formulated [56]. The problem is solved by the bipartite graph and heuristic gradient descent algorithm (HGDA). The simulation results show that the low latency is achieved and the stability among vehicles are guaranteed.

A channel selection-based resource allocation technique is studied [57]. A DRL algorithm is used to support the RSUs in making decision to select channels and allocate these channels to the vehicles on demand. In addition, the state-transition probability distribution of the cloud computing system matrices is derived by an infinite horizon semi-Markov decision process (SMDP) model. The maximum of the total long-term expected rewards of the cloud computing system is achieved by optimally allocating the resources to serve the requesting users. Besides, a semi-Markov decision process (SMDP) model-based resource allocation is proposed to immediately support service requests from VUEs [58]. In this paper, RSUs are integrated with vehicular cloud computing (VCC). Due to the nonfixed arrival time and V2V and V2I links, the vehicular network becomes heterogeneous. The reliable vehicular communication is achieved via this proposal. A DRL approach-assisted resource allocation optimization is represented [59]. The problem is considered as a joint problem of spectrum and computation power allocation in the vehicular network. Then, both single-agent RL and multiagent RL are used to find the optimal solution.

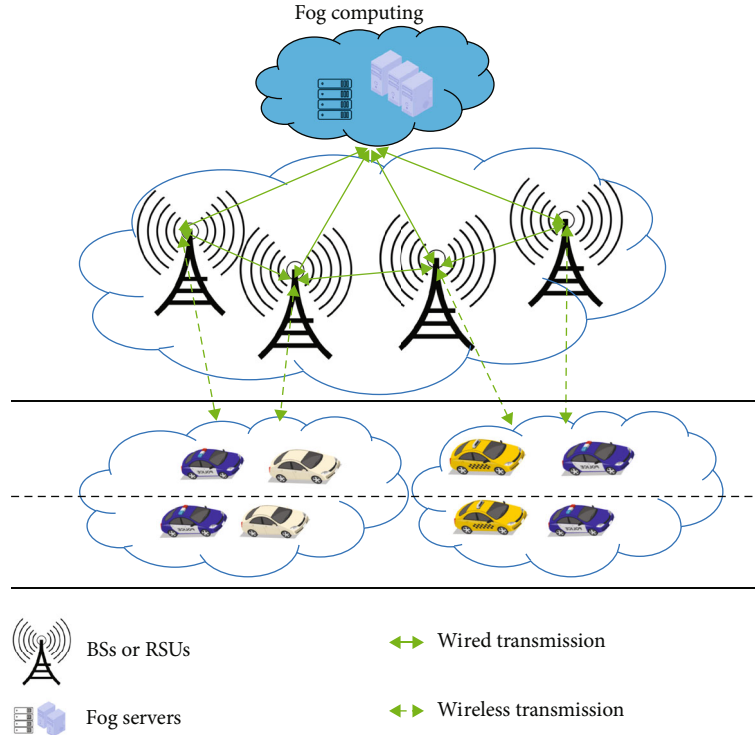


FIGURE 9: A fog computing model for the vehicular network.

A DRL approach-based resource allocation outperforms the conventional optimization problems and supervised learning approaches in the scenario of the heterogeneous vehicular network. However, the problems of high latency, security, and privacy from cloud computing cause unreliable low-latency communication. Hence, a fog computing network is proposed to improve latency, which remains in the cloud computing network for vehicles.

4.2. Computation Power and Resource Allocation in Fog Computing for the 5G Vehicular Network. A fog computing is known as a fog networking. The data processing, computation power, and resources can be offloaded from the cloud servers to the fog servers. It is developed to serve a large number of heterogeneous and distributed devices [60]. A fog computing network is illustrated in Figure 9. In this scenario, RSUs and macro-BSs are deployed on the ground to serve their vehicles. There are wireless communications among V2V and V2X links. On the other hand, the connections between RSUs and BSs and between these units to the fog computing servers are wired transmission methods. The fog servers are deployed closer to the end users (e.g., VUEs) than cloud servers. Thus, the latency due to data transmission, compared to cloud computing, is significantly reduced. Specifically, during the peak time, this approach reduces the processing delay and the responding time to end VUEs [61].

4.2.1. Optimization Theory Techniques. A joining of an optimization for user association, radio resource allocation, and power utilization is investigated [62]. In this article, a cross-computing layer, consisting of a cloud and a fog network, is designed to allocate computation power. It takes responsibility

of the global traffic management and the large-scale traffic light control. The joint optimization problem is solved by an iterative optimization algorithm. By this approach, the transmission delay is reduced. Besides, a contract-based incentive mechanism and a matching-based computation task assignment are designed [63]. The optimization problem is solved by a pricing-based stable matching algorithm. A fog-enhanced radio access network (FeRAN) is studied to improve the diverse QoS services [64]. Two resource management schemes, namely, fog resource reservation and fog resource reallocation, are proposed. The transmission latency is reduced during peak traffic time.

A non-orthogonal multiple access- (NOMA-) based fog computing vehicular (FCV) network architecture is proposed [65]. And a reinforcement learning (RL) to tackle with user mobility is also proposed [66]. The subchannel and power allocation are optimized via a chemical reaction optimization (CRO) algorithm and a real-coded chemical-reaction optimization (RCCRO) algorithm. The results show that the energy efficiency is maximized while the transmission latency is significantly reduced. An integration of user association and resource allocation is studied [67]. The joint optimization is a mixed-integer nonlinear program. A Perron-Frobenius theory is proposed to reduce transmission delay. In addition, in article [68], a distributed computation offloading and a resource allocation algorithm are integrated to support vehicular networks. Since, the joint optimization problem is nonconvex and NP-hard, a distributed computation offloading and a resource allocation algorithm are proposed to solve the collaborative computation offloading and the resource allocation optimization (CCORAO). Accordingly, the computation time is reduced and system utility is improved.

Despite, providing reduction in transmission time, these optimization problems are nonconvex and NP-hard to solve. Thus, supervised learning algorithms need to be employed to tackle the challenges.

4.2.2. Supervised Learning-Assisted Techniques. Without prior knowledge about the traffic flow, there must be a hard problem to deal with the burst increase in the number of vehicles joining the network system (e.g., during peak time). Hence, it could be an effectively utilized resource. In this scenario, traffic flows are forecasted. Then, a deep neural network-based prediction strategy (DNN-PS) is proposed to forecast the future traffic flow in the vehicular network [69]. The results of the prediction stage is used as the reference for the network system to preallocate the radio resource for the upcoming service requests in the future. However, the DNN method only has capability to forecast a short-term traffic flow. The more the network system knows about the dynamics of traffic environment, the better the resource allocation. Therefore, an LSTM approach-based time-series traffic-flow prediction is proposed [70]. By adopting the LSTM, both the short-term traffic flow forecasting and the long-term traffic flow forecasting are captured. In the practical scenario, the traffic flow is extremely dynamic. Therefore, how to obtain the training dataset for training the ML model is a big challenge. On the other hand, due to the missing data issue, the ML algorithm can not get the high-accuracy traffic flow prediction. Then, it leads to wrong resource preallocation. Not only the vehicles can not use the radio resource at the right time, but also the network system performance is not efficient.

An LSTM approach-based traffic flow prediction with missing data is proposed [71]. Specifically, a multiscale temporal smoothing is adopted to cope with the missing data. In [72], an LSTM-DNN approach is employed to predict the vehicles' movement and the parking status. This traffic information is forecasted in both a short-term resource allocation and a long-term resource allocation for vehicular fog computing (VFC). Then, the forecasted results are used as references to allocate the spectrum among vehicles via a RL decision making at the corresponding RSU. The transmission delay and computation delay are reduced by this proposal. In the work of [73], the authors propose a RL approach-based radio resource allocation algorithm to consider the future network status. The time division duplex (TDD) configuration is changed in the training phase. In addition, the action is chosen to consider the future network status. The reward for the agent is maximized based on the results of forecasting. The results outperform in throughput of the vehicular network. Besides, the packet loss is reduced.

The aforementioned supervised and unsupervised learning approaches show good performance in latency reduction. However, in the heterogeneous vehicular network, how to obtain the training data from the practical scenario faces a big challenge. Therefore, with the benefits of the DRL approach, it is taken to deal with this challenge.

4.2.3. DRL-Based Techniques. Due to the lack of training data, a DRL algorithm is proposed to deal with this challenge.

According to article [74], the wireless transmission bandwidth and the computation power contribution among vehicles are investigated. A DRL approach-based contract theory is proposed to manage these resources to reduce complexity and to avoid decision collisions among vehicles. In the same scenario, in the work of [75], a Markov decision process (MDP) model is used to present the incorporation of the computation power and the storage resource. Perception-reaction time (PRT) is the time consumption for safe driving reaction. A proposal of a DRL algorithm-based online PRT optimization is represented. The incorporation between the information-centric networking (ICN) and the fog resource virtualization is analyzed. The latency reduction is achieved via this approach. Besides that, a combination of a deep neural network (DNN) and an actor-critic (A3C) is adopted [76]. The objective is to effectively utilize computation power. Thus, the user satisfaction is achieved by latency reduction.

In another scenario, vehicles are used as fog nodes to serve mobile users [77]. In practice, vehicles are characterized by mobility. It leads to a challenge of selecting a suitable fog node to serve the users. In this article, an effective resource allocation on vehicles, based on a multiobjective optimization, is proposed to serve their own users. Then, the problem is solved by the nondominated sorting genetic algorithm. In [78], a resource decision making, based on the Markov decision process (MDP), is proposed. The approach of software-defined vehicular-based fog computing (SDV-F) is studied. From that, the computation power is maximized at the fog layer. While the end-to-end latency is significantly reduced. In the same context, a DRL algorithm is used in paper [79] to minimize task processing time at the fog servers. Further, the energy efficiency is maximized for the fog computing. With the support of the fog computing, processing time and transmission time are reduced compared to the cloud computing.

In general, a DRL approach-based technique helps to burst the computation capacity and to reduce the latency compared to the cloud computing. Even though a better performance is achieved, it still remains the latency issue due to the burst increase in the number of vehicles. Thus, an edge computing technique is proposed to achieve a better performance.

4.3. Computation Power and Resource Allocation in Edge Computing for the 5G Vehicular Network. An edge computing is a solution to offload the processing, computing, and resources from the cloud servers and the fog servers to the edge servers. The edge servers are deployed more closer to VUEs than the fog servers. Recently, BSs can be equipped with edge servers. Accordingly, the delay jitter caused by the remote cloud and fog computing is reduced [80]. Figure 10 illustrates an edge computing model for the vehicular network. In this network system, there are wireless transmission methods among V2V, V2X, and wired transmission methods among BSs. At individual BSs, there is an edge computing server. Thus, an edge computing technique provides a sustainable and low-cost solution for the vehicular network.

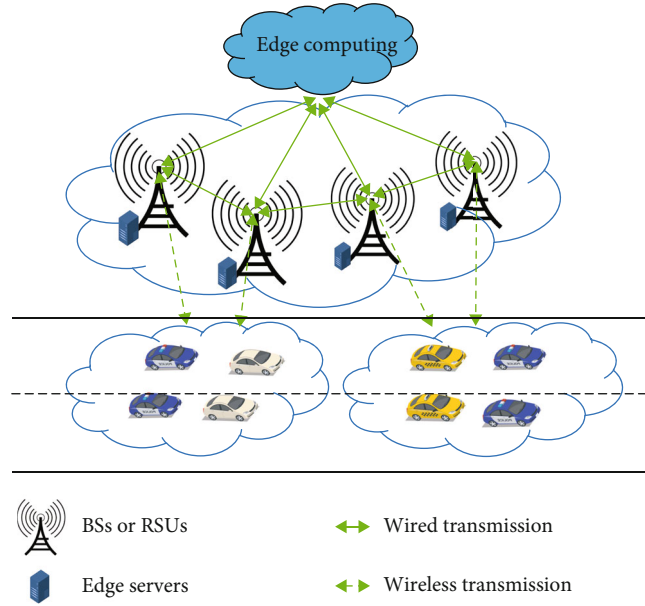


FIGURE 10: An edge computing model for the vehicular network.

4.3.1. Optimization Theory Techniques. An adaptive and online resource allocation for enhancing user experience (ARAEUE) is designed [81]. It aims at minimizing the computing loss in the vehicular edge computing network. The radio and computing resources are researched with the unknown network states. In addition, a mobility-aware greedy algorithm is studied in [82]. The amount of edge cloud resources for individual vehicles is determined by a software-defined vehicular edge computing. In this model, a controller has responsibility of task offloading strategy and edge cloud resource allocation strategy. The success probability of total task execution versus the number of vehicles, vehicular mobility, and completion time limit are achieved by the proposal. Another method, a dual-side cost of a smart vehicular terminal, is reduced [83]. The approach is about offloading decision making and allocating radio resource according to the results from offloading computation. Moreover, transmitting power and server provision are studied while the network stability is guaranteed. The results show the tradeoff between cost and queue backlog. The latency is achieved by the iterative radio resource allocation.

A uncertainty-aware resource allocation is investigated [84]. The goal is to assign arriving requests to a BS so that all the requests are served on time. It aims at maximizing service requests on time among BSs under the radio resource scarcity constraints. The simulation shows the good performance in radio resource allocation while the latency is reduced. Another approach is proposed in [85]. A computation offloading process is investigated. The minimum assignable wireless RB level in vehicular edge cloud computing (VECC) is considered to guarantee the diverse QoS requirements. In addition, the value density function is studied to measure the cost effectiveness of allocated resources and energy savings. Then, these problems are solved by theoretical discovery-based a low-complexity heuristic resource allocation algorithm. The energy efficiency is achieved by this approach.

These techniques achieve the good performances in latency reduction and energy efficiency. However, the optimization problems are nonconvex and NP-hardness. Then, they are hard to be solved. Then, an ML approach is used to tackle this challenge.

4.3.2. Supervised Learning-Assisted Techniques. A K-nearest neighbor (KNN) algorithm is used to select the task offloading platform. It includes cloud computing, mobile edge computing, and local computing. From the selected results of KNN, an RL model is applied to allocate computational resources for all VUEs. Thus, the system complexity in non-local computing is reduced. Besides that, an online RL approach (ORLA) based on a distributed user association algorithm is investigated to guarantee the QoS requirement in the vehicular network [86]. Thus, the network load is balanced while the latency is reduced according to intelligent radio resource allocation. Similarly, two game theories are studied [87] to verify the effectiveness of the load balancing scheme. The proposal of minimizing the processing delay of vehicles in computation tasks under the constraints of their maximum permissible delays is analyzed. Then, a SDN-based load balancing task offloading scheme in fiber-wireless (FiWi) techniques is designed. Thus, the computing networks are enhanced while low latency is achieved. The age of information- (AoI-) aware radio resource management problem is designed for the 5G vehicular network [88]. A decentralized online testing at the VUE pairs by an LSTM and a DRL is deployed. It enables bandwidth allocation and packet scheduling decision at RSUs. Even though only the partial network state is observed, without priori statistic knowledge of network dynamics, the effective resource utilization is achieved by this approach.

A convolutional neural network (CNN) is embedded in the DNN model [89]. It is used to approximate the offloading scheduling policy and value function. Then, a DRL

approach-based offloading scheduling method (DRLOSM) is designed to not only optimize the number of retransmitted tasks and cost but also reduce energy consumption. An alternating direction method of multipliers (ADMM) is investigated [90]. It is a distributed algorithm. An information-centric heterogeneous network framework is designed to enable content caching and computing. The designed network system allows sharing communication, computing, and caching resources among users with heterogeneous virtual services. Similarly, in the work of [91], a mobility-aware task offloading, processed by VEC servers, deployed at the access points is investigated. If there is overload at one access point, this access point will share the collected overloading tasks to the adjacent server at the next access point. It is implemented by a cooperative MEC server scenario on the vehicle's moving direction. Thus, the computing and processing delays are reduced while enhancing the vehicular network.

To deal with the hardness in obtaining the training dataset for supervised learning models, a DRL approach-based techniques is assigned to cope with the challenge.

4.3.3. DRL-Based Techniques. A Q-learning is studied to allocate transmit power, subchannels, and computing resources in edge computing [92] for the vehicular network. More specifically, a SDN-assisted edge computing is designed to reduce the signaling overhead. The results show the reduction in system overhead, network load, and spectrum utilization. Similarly, a DRL algorithm is applied to participate in tasks and to select computing servers for vehicles in a collaborative MEC computing [93]. Specifically, a model is designed with MEC servers at both macro-BSs and at edge nodes. The low-latency reliable services are achieved in the vehicular network by this approach. Besides that, an offline training of a RL and deep deterministic policy gradient (DDPG) is proposed in the work of [94]. The model learns the network dynamics and then allocates spectrum, computation power, and storage resource among vehicles accordingly. With the same approach, an MEC migration framework is studied to reduce computation and communication load of a MEC server to other MEC servers due to overload in processing [95]. The computation migration is done by an adaptive learning technique with a DRL algorithm. These aforementioned approaches guarantee ultra-reliable low-latency communication.

Recently, UAV-assisted communication is an emerging topic. The scenario is illustrated in Figure 11. In this approach, multiaccess edge computing servers are mounted in both BSs, RSUs, and UAVs. UAVs act as flying BSs in order to provide connections in the case of broken links among vehicles. A multi-agent DDPG-based resource allocation in multi-access edge computing servers is proposed in the work of [96]. The same as [96], in [97], multiaccess edge computing servers are equipped in both macro-eNodeBs and UAVs. An offline learning approach with the RL algorithm and DDPG is studied. The optimal latency is achieved by managing the spectrum, computation, and caching resources. Besides that, the scenarios of unstable energy arrival, stochastic computation tasks from VUEs, and time-varying channel state are analyzed

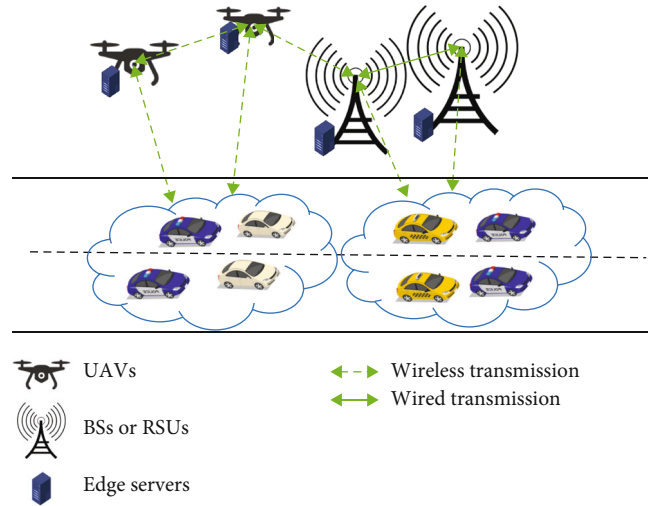


FIGURE 11: UAV-assisted vehicular communication model.

to reduce latency [98]. UAVs are dynamically deployed to support the computation offloading. Similarly, in the work of [99], a DRL-based optimal channel selection is proposed. The objective is to train the network according to the historical data. Then, it finds the optimal available CR channel in CR-VANET.

Among the three computing layers (i.e., cloud computing, fog computing, and edge computing), an edge computing network supported by AI techniques, especially DRL algorithms, is considered as the best approach to reduce latency and to increase energy efficiency in the 5G vehicular network. In addition, UAV-assisted communication for multiaccess edge computing is a promising approach to achieve the diverse QoS requirements in the 5G and 6G vehicular networks.

5. FDRL-Based UAV-Assisted Vehicular Communication

5.1. UAV-Assisted Vehicular Network. A drone, known as an unmanned aerial vehicle (UAV), is considered to be an important element in the Internet of things (IoT) and wireless communication in 5G and toward 6G vehicular networks. UAVs, equipped with communication devices, are dynamically deployed to support broadband wireless communication [100]. In a communication scenario, UAVs are deployed as flying BSs to provide wireless access points to users. In the traditional vehicular network, including vehicles, BSs, and RSUs, messages are frequently blocked by neighbor vehicles and buildings. The reliable links are provided by adding UAVs to the traditional vehicular network; messages can be broadcasted via line-of-sight (LOS).

Recently, UAVs act as relays to assist vehicular communication on the highway in case of disaster situations (e.g., flood and earthquake) [18]. In this scenario, BSs and RSUs are damaged and the communication links are broken. Thus, UAVs with the limited battery are carefully designed to maximize the minimum average range for individual vehicles by optimizing the UAV trajectory and radio resource allocation.

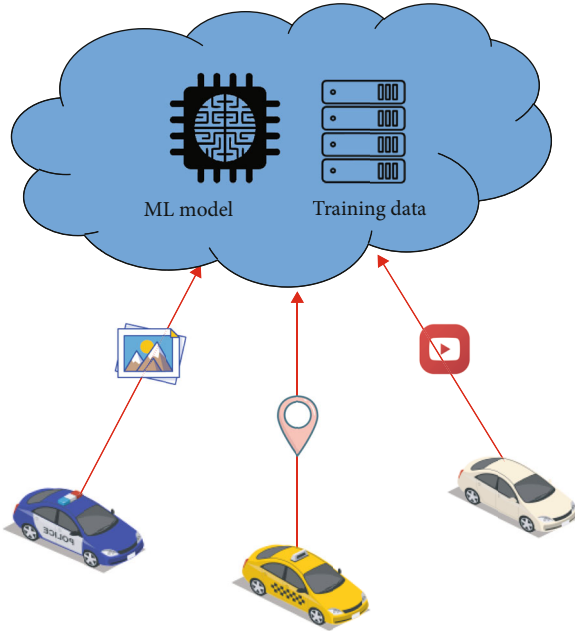


FIGURE 12: A centralized machine learning model.

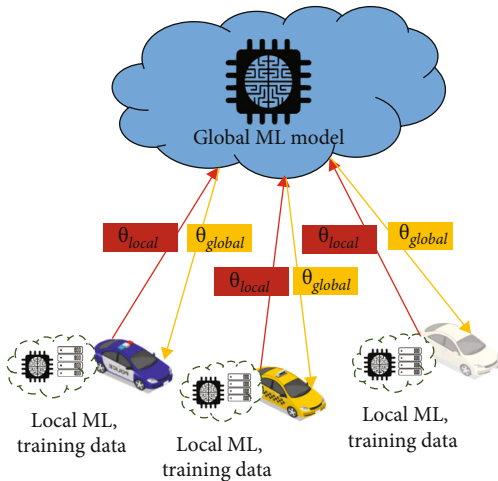


FIGURE 13: A distributed federated learning model.

Similarly, UAVs are dynamically deployed against jamming in transportation [101]. In the scenario, vehicles are equipped with sensors such as cameras and GPS devices. The vehicles gather the sensing information and sends messages to the servers via the serving RSUs [102]. Both UAVs and RSUs received the messages from the vehicles and then the UAVs decide whether to connect to the servers via the RSUs or not. A hotbooting policy hill climbing-based UAV relay strategy is proposed to help the vehicular network resist jamming. The bit error rate of the vehicular message is reduced and the utility is increased.

An energy-aware dynamic power optimization problem is formulated under the constraint of the evolution law of the energy consumption state for each vehicle [103]. In this scenario, both cooperation and noncooperation among vehicles are investigated to obtain the optimal dynamic power

allocation of the vehicles with a fixed UAV trajectory [104]. Similarly, a deep deterministic policy gradient- (DDPG-) based resource management in UAVs and macro-eNodeBs (MeNBs) is investigated [97]. In this work, MEC servers are mounted into MeNBs and UAVs to cooperatively make association decisions and allocate the proper amount of resources to vehicles. A DDPG-based spectrum, computing, and caching resource allocation for vehicles are studied while the MEC servers acting as learning agents in a RL algorithm. As an improvement of [97], the work in [96] proposes a multi-agent RL deep deterministic policy gradient (MADDPG) to manage radio resource and computing power. These approaches show the good performance in reducing latency and enhancing network connection.

5.2. Federated Deep Reinforcement Learning-Based Vehicular Network

5.2.1. Federated Learning Concept.

In 6G, intelligent radio (IR) and self-learning with proactive exploration, such as distributed federated learning (FL), will be constructed, because it can deal with the FemBB, EDURLLC, ELPC, and ELPC QoS requirements in the 6G vehicular network scenario [105]. A comparison between centralized learning and federated learning is illustrated in Figures 12 and 13, respectively. On one hand, in a centralized machine learning model, an ML algorithm is employed at the server. All the raw data from VUEs are transmitted to their servers for data processing, computing, and so on. On the other hand, in a distributed federated learning, an ML algorithm is employed at the server and the ML algorithms are also employed at VUEs. Only parameters of the local ML algorithms at VUEs are transmitted to the server. Then, the privacy is guaranteed and the latency is reduced, compared to a centralize machine learning model.

More specifically, FL is a distributed machine learning technique where multiple vehicles train a common model by using their own local dataset. The vehicles send the update parameters of the common model instead of all raw data to the central server. According to the training method, the FL can enrich result without sacrificing the privacy of the vehicle data. The basic steps for a FL technique are as follows:

- (1) The central server initializes random parameters or trains an ML algorithm with its own dataset. From the training, the global parameters of the ML algorithm are achieved at the central server
- (2) In the end vehicle selection, the central server identifies end vehicles (e.g., VUEs) to be involved in the model training
- (3) In global parameter distribution, after identifying the end vehicles, the global parameters are distributed from the central server to the selected end vehicles
- (4) In the local training model, the end vehicles train their own ML algorithms with the received parameters and their own generated dataset to calculate local parameters

- (5) In local parameter feedback, the local parameters are sent from the end vehicles to the central server
- (6) In parameter aggregation, the central server aggregates all the local parameters from its end vehicles via an algorithm such as a federated averaging (FeAvg) method
- (7) In global model testing, the aggregated global parameters are used to test with the dataset at the central server. After that, the new global parameters are calculated by the central server
- (8) In new global parameter distribution, the central server distributes the new global parameters to the end vehicles for local training. The training process is iterative until the global model at the central server finds the desirable parameters

5.2.2. Federated Deep Reinforcement Learning. A federated deep reinforcement learning (FDRL) technique is based on a DRL model to train collaboratively learning models on the end vehicles while reducing communications overhead and guaranteeing privacy. A DRL algorithm with agent α is given:

$$Q_\alpha^*(s, a, \theta_\alpha) \leftarrow Q_\alpha(s, a, \theta_\alpha) + \alpha \left[r + \gamma \max_{\{a'\}} Q_\alpha \cdot (s', a', \theta'_\alpha) - Q_\alpha(s, a, \theta_\alpha) \right], \quad (9)$$

with the loss function

$$L(s, a, \theta_\alpha) = \left(r + \gamma \max_{\{a'\}} Q_\alpha(s', a', \theta'_\alpha) - Q_\alpha(s, a, \theta_\alpha) \right)^2. \quad (10)$$

Once θ_α is learnt, the policy $\pi^* = \arg \max_{a \in A} Q(s, a, \theta_\alpha)$ is achieved. Then, the FDRL problem is defined by given transitions (memory) $\mathcal{D}_\alpha = \{s_\alpha, a_\alpha, s'_\alpha, r_\alpha\}$ collected by agent α . It aims at federatively building policies π^* for agent α . Note that states, actions, Q-functions, and policies with respect to agent α by $s_\alpha \in S_\alpha$, $a_\alpha \in A_\alpha$, Q_α , and π_α^* , respectively. Thus, a new deep Q-network, namely, federated deep Q-network is denoted by Q_f

$$Q_f(\theta_\alpha, \theta_f) = (Q_\alpha(s, a, \theta_\alpha), \theta_f). \quad (11)$$

The state of FDRL is defined as $S_{\alpha,n}(t) = \{Z_{\alpha,n}(t)\}$ where $Z_{\alpha,n}(t)$ represents a set of request services from vehicles $\mathcal{N} = (1, 2, \dots, n, \dots, N)$. At the beginning of the decision period $t = 0$, all the N vehicles receive the global parameters $\theta_f(t)$ responding to Q_f from the central server. After that, the vehicles compute their own local models $\theta_{\alpha,n}(t)$ with the received parameters from the center server, based on the current service request $Z_{\alpha,n}(t)$ by their own training dataset. Then, each vehicle sends its computed parameter $\theta_{\alpha,n}(t+1)$ to the central server. The central server aggregates these updates and averages them via an algorithm (e.g., federated averaging

algorithm) to get a new global parameter $\theta_f(t+1)$. Finally, the center server distributes $\theta_f(t+1)$ to its vehicles. The detail training procedure can be seen in Algorithm 1.

5.3. Federated Deep Reinforcement Learning-Based UAV-Assisted Vehicular Network. In order to achieve the ultra-reliable low-latency vehicular communication in a highly mobile multi-access vehicular environment, wireless network resources must be effectively utilized. On one hand, computation and processing delay must be reduced to support URLLC for the vehicular network. On the other hand, data privacy should be guaranteed in the vehicular network to support safety application. Thus, a FL technique can be easily integrated with edge computing to reduce computation, processing latency, and privacy for the end vehicles [106]; a joint power and radio resource allocation with the support of a FL technique integrated with a MEC scheme is proposed to achieve ultra-reliable low-latency vehicular communication. The FL is used to estimate the tail distribution of the queue length in the network. From the estimated results, power and the radio resource are allocated to the end vehicles. The results show that the latency and communication overhead are reduced, compared to other centralized learning schemes.

The problem of how to obtain training data from the dynamic environment of the vehicular network is a big challenge. In addition, in order to reduce latency and communication overhead while guaranteeing privacy, FDRL is proposed. Due to the broken communication links in fast mobility and disaster, UAVs, equipped with communication devices and other electronic devices, are used to provide connectivities. Recently, there are existing techniques of the UAV-assisted vehicular network as the aforementioned analyses. Based on DRL algorithms, the UAV networks can solve the issues not only in obtaining data but also in navigating the UAVs smoothly in the missions. Further, one of the most stringent requirements in the vehicular network is ultra-reliable low-latency communication. Thus, FRDL is applied to deal with both obtaining training data challenge, reducing latency, and guaranteeing privacy.

6. Challenges, Open Issues, and Future Direction

6.1. A Multiaccess Edge Computing in SDN and NFV. In 5G NR, the vehicular ad-hoc network has characteristics of a heterogeneous and large-scale structure. These features make it difficult to efficiently deploy ML algorithms [107]. Recently, for the 5G vehicular network, SDN and NFV are proposed to enable software network functions and networking slicing. It can cope with the heterogeneous networks and the diverse QoS services. On one hand, SDN and NFV are technologies developed to achieve the deserve QoS requirements (i.e., URLLC, eMBB, and mMTC) in the 5G NR. On the other hand, to deal with the requirement in computation capacity, resource allocation, and storage resource, a multiaccess edge computing is proposed, which is illustrated in Figure 10. It can achieve user satisfaction by guaranteeing diverse QoS requirements in the 5G vehicular network [108]. Moreover, the 6G vehicular network is with ultralow delay and super-

Input: state space $S_\alpha(t)$, action space $A_\alpha(t)$, reward R
Output: θ_α, θ_f
Initialization:
1: Central server:
 Initialize the global DRL model Q_f with random value $\theta_f(0)$ at the beginning of decision period $t = 0$
2: Local vehicles:
 Initialize the local DRL models $Q_{\alpha,n}$ with values $\theta_{\alpha,n}(0), n = (1, 2, \dots, N)$
 Download $\theta_f(0)$ from the central server and let $\theta_{\alpha,n}(0) = \theta_f(0), n = (1, 2, \dots, N)$
 : Initialize replay memory \mathcal{D}_α
Iteration: For each decision period $t = 0$ to T do
3: **function** FL($Z_{\alpha,n}(t)$)
 Local vehicles:
4: **while** $t > 0$ **do**
5: **for** each vehicle $n \in \mathcal{N}$ in parallel **do**
6: download $\theta_f(t)$ from the controller
7: let $\theta_\alpha(t) = \theta_f(t)$
8: train the DRL agent locally with $\theta_{\alpha,n}(t)$ on the current service requests $Q_n(t)$
9: upload the trained weights $\theta_{\alpha,n}(t+1)$ to the central server;
10: observe $\theta_{\alpha,n}(t+1), n = (1, 2, \dots, N)$ in \mathcal{D}_α
10: **end for**
11: central server:
12: receive all weights $\theta_{\alpha,n}(t)$ updates;
13: perform federated averaging;
14: broadcast averaged weights $\theta_f(t+1)$
15: **end while**
16: **end function**

ALGORITHM 1: FDRL.

high network capacity while the resources are limited. Therefore, the combination between a softwareized network function and a multiaccess edge computing is needed to be investigated. The approach allows multiplexing different data traffic patterns. Thus, it should be more investigated to satisfy the diverse QoS requirements over 5G and toward 6G vehicular networks.

6.2. FDRL-Based UAV-Assisted Vehicular Network

6.2.1. FDRL for the Vehicular Network

(1) *FDRL-Based Resource Allocation.* Conventional mathematical optimization approaches face challenges of the complex computation and the complex vehicular environment. In addition, due to the dynamic environment in the vehicular network, how to obtain dataset for training machine learning models is a big issue. Therefore, a DRL approach is used and trained to cope with the lack of dataset. However, to reduce latency and increase privacy demands, a FL technique that uses a DRL algorithm in the local training models at the end vehicles is expected to be a promising technique to solve the power and radio resource allocation problem for 5G and toward 6G vehicular networks.

(2) *Communication, Computing, and Caching Strategies for FL.* In order to reduce the computation and processing delay for the end vehicles, an FL approach is integrated with MEC servers, which deployed the corresponding

BSSs, and RSUs are proposed. More specifically, DRL algorithms, which are employed in the MEC servers and in the end vehicles, are trained for a specific purpose. The communication, computing, and caching strategies for the FL model needs to be carefully considered to achieve a better efficient network performance while guaranteeing the heterogeneous QoS requirements.

(3) *Collaborative Intelligence.* The vehicular network becomes heterogeneous, including vehicles, RSUs, BSs, UAVs, edge servers and other devices. A FL technique-based efficient resource allocation strategy among these heterogeneous devices needs to be investigated to ensure the diverse QoS requirements in 5G and toward 6G vehicular networks.

6.2.2. FDRL for UAVs

(1) *UAV Dynamic Deployment.* One of the most important characteristics in the vehicular network is a quite high mobility. In addition, the deployment of UAVs mostly depends on users' behavior. A centralized learning algorithm can help to collect and predict the users' behavior. However, it is difficult. Since the central servers at the UAVs require different data from different vehicles and due to their resource limitation, thus, FDRL can be used to cope with these challenges. In this approach, vehicles can learn by themselves with their own dataset and report their behaviors to the central servers such as MEC servers deployed in UAVs.

(2) *Spectrum Allocation*. UAVs act as flying BSs and must assign spectrum resource for VUEs. It forces UAVs to be dynamically adapted and to serve real-time services. A centralized learning method may introduce more latency. Thus, a FDRL technique should be used to dynamically assign the spectrum resource for VUEs in a real-time scenario. This is because UAVs can collaboratively generate a prediction model regarding to the historical spectrum allocation data.

(3) *UAV Capacity*. Due to the limitation in computation, storage, processing resources, and energy limitation of the battery, UAV trajectory planning must be carefully designed. Moreover, within communication among UAVs (i.e., air-to-air communication), it requires to share UAVs' mobility and energy level to each other to more effectively utilize resources to serve all VUEs. However, due to the privacy concern, a centralized learning approach can be replaced by a decentralized learning approach (e.g., FDRL) to learn local energy consumption and predict the future energy consumption. Thus, it allows UAVs to determine their trajectories.

6.2.3. *FDRL-Assisted UAV-Based Vehicular Network*. In order to reduce latency due to computation and processing missions, MEC servers are mounted in both macro-BSs, RSUs, and UAVs. Thus, there are 5 main features needed to satisfy when using a distributed FDRL in vehicular communication. Firstly, the data gathered by the edge devices (e.g., VUEs) cannot be transmitted to the cloud, fog, and edge servers. Secondly, the training model must be fast due to the exchanges of parameters between the global model and the local models (e.g., in VUEs). Thirdly, the exchange parameters between the global model and the local models must be fast. Fourthly, the gathered data from the edge devices are needed to be labeled fast and accurately on the same devices. Lastly, the edge devices must have enough computation capacity and storing resources to effectively train the local data models.

7. Conclusions

In this paper, the recent advanced techniques of the conventional optimization theory, ML, especially DRL-based resource managements, are reviewed. These techniques are discussed in cloud, fog, and edge layers to guarantee diverse QoS requirements in the 5G vehicular network. Besides, a distributed FL technique, especially a FDRL approach-based UAV-assisted vehicular network, to provide connectivity while reducing latency, is investigated. In addition, a mechanism of FDRL-based vehicular communication is proposed. Then, the challenges, open issues, and future direction in 5G and toward vehicular networks are represented. Firstly, the analysis of a multiaccess edge computing, combined with a software-ization approach, is discussed to point out future research direction. Then, the challenges and future research direction for distributed a FDRL technique-based UAV-assisted 5G and toward 6G vehicular communication are analyzed and provided. The three main open research directions are addressed, including a FDRL technique-based vehicular network, a FDRL technique-based UAVs, and a

FDRL technique-based UAV-assisted vehicular communication. More specifically, the problems of scheduling, computation power, transmitting power and radio resource management, and latency reduction should be addressed when using a FDRL technique-based UAV-assisted vehicular communication. Secondly, the learning method issues consist of data diversity, labeling, and efficient model training. Finally, the computing capacity of devices needs to be defined. In general, all the aforementioned communication challenges can be taken by ML algorithms, especially DRL and FDRL approaches. Accordingly, the diverse QoS requirements can be guaranteed in 5G, toward 6G vehicular networks.

Data Availability

All the data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Thai Nguyen University of Technology (TNUT), Ministry of Education and Training (MOET), Viet Nam, for the support.

References

- [1] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A survey on 5g usage scenarios and traffic models," *IEEE Communications Surveys Tutorials*, vol. 22, no. 2, pp. 905–929, 2020.
- [2] S. A. A. Balamurugan, J. F. Lilian, and S. Sasikala, "The future of India creeping up in building a smart city: intelligent traffic analysis platform," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 518–523, Coimbatore, India, 2018.
- [3] T. Huang, W. Yang, J. Wu, J. Ma, X. Zhang, and D. Zhang, "A survey on green 6g network: architecture and technologies," *IEEE Access*, vol. 7, pp. 175758–175768, 2019.
- [4] L. Liang, H. Peng, G. Y. Li, and X. Shen, "Vehicular communications: a physical layer perspective," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10647–10659, 2017.
- [5] A. Aslan, G. Bal, and C. Toker, "Dynamic resource management in next generation networks based on deep q learning," in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, Gaziantep, Turkey, 2020.
- [6] M. Series, "Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond," *Recommendation ITU*, vol. 2083, 2015.
- [7] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. J. A. Zhang, "The roadmap to 6g: Ai empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, 2019.
- [8] M. Latva-aho, K. Leppänen, F. Clazzer, and A. Munari, *Key Drivers and Research Challenges for 6g Ubiquitous Wireless Intelligence*, 6G Flagship, University of Oulu, 2020.
- [9] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6g wireless communication systems: applications,

- requirements, technologies, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, 2020.
- [10] T. Nakamura, "5g evolution and 6g," in *2020 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*, p. 1, Hsinchu, Taiwan, 2020.
- [11] F. Tariq, M. R. A. Khandaker, K. K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6g," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 118–125, 2020.
- [12] S. Zhang, C. Xiang, and S. Xu, "6g: connecting everything by 1000 times price reduction," *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 107–115, 2020.
- [13] Y. Sun, J. Liu, J. Wang, Y. Cao, and N. Kato, "When machine learning meets privacy in 6g: a survey," *IEEE Communications Surveys Tutorials*, vol. 22, no. 4, pp. 2694–2724, 2020.
- [14] H. Viswanathan and P. E. Mogensen, "Communications in the 6g era," *IEEE Access*, vol. 8, pp. 57063–57074, 2020.
- [15] M. Zamzam, T. El-Shabrawy, and M. Ashour, "Game theory for computation offloading and resource allocation in edge computing: a survey," in *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, pp. 47–53, Giza, Egypt, 2020.
- [16] K. R. Dalal, "Analysing the role of supervised and unsupervised machine learning in iot," in *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 75–79, Coimbatore, India, 2020.
- [17] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, pp. 1273–1282, Ft. Lauderdale, FL, USA, 2017.
- [18] M. Samir, M. Chraiti, C. Assi, and A. Ghayeb, "Joint optimization of UAV trajectory and radio resource allocation for drive-thru vehicular networks," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, Marrakesh, Morocco, 2019.
- [19] A. Sharma and L. K. Awasthi, "A comparative survey on information dissemination in heterogeneous vehicular communication networks," in *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pp. 556–560, Jalandhar, India, 2018.
- [20] K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, and Y. Zhou, "Heterogeneous vehicular networking: a survey on architecture, challenges, and solutions," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2377–2396, 2015.
- [21] A. M. Căilean and M. Dimian, "Current challenges for visible light communications usage in vehicle applications: a survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2681–2703, 2017.
- [22] W. M. Danquah and D. T. Altılar, "Vehicular cloud resource management, issues and challenges: a survey," *IEEE Access*, vol. 8, pp. 180587–180607, 2020.
- [23] M. Noor-A-Rahim, Z. Liu, H. Lee, G. M. N. Ali, D. Pesch, and P. Xiao, "A survey on resource allocation in vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, 2020.
- [24] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, "Securing connected & autonomous vehicles: challenges posed by adversarial machine learning and the way forward," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 998–1026, 2020.
- [25] F. Tang, Y. Kawamoto, N. Kato, and J. Liu, "Future intelligent and secure vehicular network toward 6g: machine-learning approaches," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 292–307, 2020.
- [26] M. A. Hossain, R. M. Noor, K. L. A. Yau, S. R. Azzuhri, M. R. Z'aba, and I. Ahmedy, "Comprehensive survey of machine learning approaches in cognitive radio-based vehicular ad hoc networks," *IEEE Access*, vol. 8, pp. 78054–78108, 2020.
- [27] Y. Zeng, G. Wang, and B. Xu, "A basal ganglia network centric reinforcement learning model and its application in unmanned aerial vehicle," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 2, pp. 290–303, 2018.
- [28] S. Chen, M. Wang, W. Song, Y. Yang, Y. Li, and M. Fu, "Stabilization approaches for reinforcement learning-based end-to-end autonomous driving," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4740–4750, 2020.
- [29] G. Chen, Y. Zhan, G. Sheng, L. Xiao, and Y. Wang, "Reinforcement learning-based sensor access control for wbans," *IEEE Access*, vol. 7, pp. 8483–8494, 2018.
- [30] J. J. Yeh, T. T. Kuo, W. Chen, and S. D. Lin, "Minimizing expected loss for risk-avoiding reinforcement learning," in *2014 International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 11–17, Shanghai, China, 2014.
- [31] A. Xanthopoulos, A. Kiatipis, D. E. Koulouriotis, and S. Stieger, "Reinforcement learning-based and parametric production-maintenance control policies for a deteriorating manufacturing system," *IEEE Access*, vol. 6, pp. 576–588, 2017.
- [32] M. Peters, W. Ketter, M. Saar-Tsechansky, and J. Collins, "A reinforcement learning approach to autonomous decision-making in smart electricity markets," *Machine Learning*, vol. 92, no. 1, pp. 5–39, 2013.
- [33] M. Hausknecht and P. Stone, *Deep Recurrent Q-Learning for Partially Observable MDPs*, 2015 AAAI Fall Symposium Series, 2015.
- [34] B. Kisacanin, "Deep learning for autonomous vehicles," in *2017 IEEE 47th International Symposium on Multiple-Valued r (ISMVL)*, p. 142, Novi Sad, Serbia, 2017.
- [35] R. I. Rasel, N. Sultana, and N. Hasan, "Financial instability analysis using ann and feature selection technique: application to stock market price prediction," in *2016 International Conference on Innovations in Science, Engineering and Technology (ICISSET)*, pp. 1–4, Dhaka, Bangladesh, 2016.
- [36] Z. Ācenal, "Smart farming becomes even smarter with deep learning—a bibliographical analysis," *IEEE Access*, vol. 8, pp. 105587–105609, 2020.
- [37] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040–53065, 2019.
- [38] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: a comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.
- [39] M. Giordani, A. Zanella, and M. Zorzi, "Millimeter wave communication in vehicular networks: challenges and opportunities," in *2017 6th International Conference on Modern Circuits and Systems Technologies (MOCASST)*, pp. 1–6, Thessaloniki, Greece, 2017.
- [40] M. Gerla, "Vehicular cloud computing," in *2012 The 11th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, pp. 152–155, Ayia Napa, Cyprus, 2012.

- [41] Y. Xiao and C. Zhu, "Vehicular fog computing: vision and challenges," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 6–9, Kona, HI, USA, 2017.
- [42] S. Buda, S. Guleng, C. Wu, J. Zhang, K. A. Yau, and Y. Ji, "Collaborative vehicular edge computing towards greener its," *IEEE Access*, vol. 8, pp. 63935–63944, 2020.
- [43] Y. Cai, F. R. Yu, and S. Bu, "Cloud radio access networks (c-ran) in mobile cloud computing systems," in *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 369–374, Toronto, ON, Canada, 2014.
- [44] A. Aliyu, A. Abdullah, O. Kaiwartya et al., "Cloud computing in vanets: architecture, taxonomy, and challenges," *IETE Technical Review*, vol. 35, no. 5, pp. 523–547, 2018.
- [45] Y. Cai, Q. Zhang, and Z. Feng, "Qos-guaranteed radio resource scheduling in 5g v2x heterogeneous systems," in *2019 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, Wai-koloa, HI, USA, 2019.
- [46] C. Guo, L. Liang, and G. Y. Li, "Resource allocation for low-latency vehicular communications with packet retransmission," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Abu Dhabi, United Arab Emirates, 2018.
- [47] L. Ding, Y. Wang, P. Wu, and J. Zhang, "Position-based user-centric radio resource management in 5g udn for ultra-reliable and low-latency vehicular communications," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Shanghai, China, 2019.
- [48] T. Sahin and M. Boban, "Radio resource allocation for reliable out-of-coverage v2v communications," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, Porto, Portugal, 2018.
- [49] Z. Lyu, Y. Wang, M. Liu, and Y. Chen, "Service-driven resource management in vehicular networks based on deep reinforcement learning," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–6, London, UK, 2020.
- [50] C. Sun, C. She, C. Yang, T. Q. S. Quek, Y. Li, and B. Vucetic, "Optimizing resource allocation in the short blocklength regime for ultra-reliable and low-latency communications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 402–415, 2019.
- [51] S. Khan, H. A. Khattak, A. Almogren et al., "5g vehicular network resource management for improving radio access through machine learning," *IEEE Access*, vol. 8, pp. 6792–6800, 2020.
- [52] J. Ye and Y. Zhang, "Pricing-based resource allocation in virtualized cloud radio access networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 7096–7107, 2019.
- [53] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, "Distributed federated learning for ultra-reliable low-latency vehicular communications," *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 1146–1159, 2020.
- [54] M. K. Abdel-Aziz, S. Samarakoon, M. Bennis, and W. Saad, "Ultra-reliable and low-latency vehicular communication: an active learning approach," *IEEE Communications Letters*, vol. 24, no. 2, pp. 367–370, 2020.
- [55] X. Chen, C. Wu, T. Chen et al., "Age of information aware radio resource management in vehicular networks: a proactive deep reinforcement learning perspective," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2268–2281, 2020.
- [56] J. Mei, K. Zheng, L. Zhao, L. Lei, and X. Wang, "Joint radio resource allocation and control for vehicle platooning in lte-v2v network," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12218–12230, 2018.
- [57] R. Pal, N. Gupta, A. Prakash, R. Tripathi, and J. J. P. C. Rodrigues, "Deep reinforcement learning based optimal channel selection for cognitive radio vehicular ad-hoc network," *IET Communications*, vol. 14, no. 19, pp. 3464–3471, 2020.
- [58] C. Lin, D. Deng, and C. Yao, "Resource allocation in vehicular cloud computing systems with heterogeneous vehicles and roadside units," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3692–3700, 2018.
- [59] L. Liang, H. Ye, G. Yu, and G. Y. Li, "Deep-learning-based wireless resource allocation with application to vehicular networks," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 341–356, 2020.
- [60] C. Tang, X. Wei, C. Zhu, Y. Wang, and W. Jia, "Mobile vehicles as fog nodes for latency optimization in smart cities," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9364–9375, 2020.
- [61] Y. Qiu, H. Zhang, K. Long, H. Sun, X. Li, and V. C. M. Leung, "Improving handover of 5g networks by network function virtualization and fog computing," in *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–5, Qingdao, China, 2017.
- [62] K. Zhang, M. Peng, and Y. Sun, "Delay-optimized resource allocation in Fog-Based vehicular networks," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1347–1357, 2021.
- [63] Z. Zhou, P. Liu, J. Feng, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Computation resource allocation and task assignment optimization in vehicular fog computing: a contract-matching approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3113–3125, 2019.
- [64] J. Li, C. Natalino, D. P. Van, L. Wosinska, and J. Chen, "Resource management in fog-enhanced radio access network to support real-time vehicular services," in *2017 IEEE 1st International Conference on Fog and Edge Computing (ICFEC)*, pp. 68–74, Madrid, Spain, 2017.
- [65] Y. Liu, H. Zhang, K. Long, H. Zhou, and V. C. M. Leung, "Fog computing vehicular network resource management based on chemical reaction optimization," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1770–1781, 2021.
- [66] A. Masaracchia, M. Nguyen, and A. Kortun, "User mobility into NOMA assisted communication: analysis and a reinforcement learning with neural network based approach," *EAI Industrial Networks and Intelligent Systems*, vol. 7, no. 25, 2021.
- [67] A. Masaracchia, D. B. Da Costa, T. Q. Duong, M. N. Nguyen, and M. T. Nguyen, "A PSO-based approach for user-pairing schemes in NOMA systems: Theory and applications," *IEEE Access*, vol. 7, pp. 90550–90564, 2019.
- [68] J. Zhao, Q. Li, Y. Gong, and K. Zhang, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7944–7956, 2019.
- [69] A. Yu, H. Yang, W. Bai, L. He, H. Xiao, and J. Zhang, "Leveraging deep learning to achieve efficient resource allocation with traffic evaluation in datacenter optical networks," in

- 2018 *Optical Fiber Communications Conference and Exposition (OFC)*, pp. 1–3, San Diego, CA, USA, 2018.
- [70] B. Yang, S. Sun, J. Li, X. Lin, and Y. Tian, “Traffic flow prediction using LSTM with feature enhancement,” *Neurocomputing*, vol. 332, pp. 320–327, 2019.
- [71] Y. Tian, K. Zhang, J. Li, X. Lin, and B. Yang, “LSTM-based traffic flow prediction with missing data,” *Neurocomputing*, vol. 318, pp. 297–305, 2018.
- [72] S. S. Lee and S. Lee, “Resource allocation for vehicular fog computing using reinforcement learning combined with heuristic information,” *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10450–10464, 2020.
- [73] Y. Zhou, F. Tang, Y. Kawamoto, and N. Kato, “Reinforcement learning-based radio resource control in 5g vehicular network,” *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 611–614, 2020.
- [74] J. Zhao, M. Kong, Q. Li, and X. Sun, “Contract-based computing resource management via deep reinforcement learning in vehicular fog computing,” *IEEE Access*, vol. 8, pp. 3319–3329, 2020.
- [75] X. Chen, S. Leng, K. Zhang, and K. Xiong, “A machine-learning based time constrained resource allocation scheme for vehicular fog computing,” *China Communications*, vol. 16, no. 11, pp. 29–41, 2019.
- [76] S. Lee and S. Lee, “Poster abstract: deep reinforcement learning-based resource allocation in vehicular fog computing,” in *IEEE INFOCOM 2019- IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs)*, pp. 1029–1030, Paris, France, 2019.
- [77] T. Mekki, R. Jmal, L. Chaari, I. Jabri, and A. Rachedi, “Vehicular fog resource allocation scheme: a multi-objective optimization based approach,” in *2020 IEEE 17th Annual Consumer Communications Networking Conference (CCNC)*, pp. 1–6, Las Vegas, NV, USA, 2020.
- [78] M. Ibrar, A. Akbar, R. Jan et al., “Artnet: Ai-based resource allocation and task offloading in a reconfigurable internet of vehicular networks,” *IEEE Transactions on Network Science and Engineering*, p. 1, 2020.
- [79] D. Lan, A. Taherkordi, F. Eliassen, and L. Liu, “Deep reinforcement learning for computation offloading and caching in fog-based vehicular networks,” in *2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pp. 622–630, Delhi, India, 2020.
- [80] X. Xu, Y. Xue, X. Li, L. Qi, and S. Wan, “A computation offloading method for edge computing with vehicle-to-everything,” *IEEE Access*, vol. 7, pp. 131068–131077, 2019.
- [81] X. Sun, J. Zhao, X. Ma, and Q. Li, “Enhancing the user experience in vehicular edge computing networks: an adaptive resource allocation approach,” *IEEE Access*, vol. 7, pp. 161074–161087, 2019.
- [82] S. Choo, J. Kim, and S. Pack, “Optimal task offloading and resource allocation in software-defined vehicular edge computing,” in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 251–256, Jeju, Korea (South), 2018.
- [83] H. Zhang, Z. Wang, and K. Liu, “V2x offloading and resource allocation in sdn-assisted mec-based vehicular networks,” *China Communications*, vol. 17, no. 5, pp. 266–283, 2020.
- [84] A. Kovalenko, R. F. Hussain, O. Semiari, and M. A. Salehi, “Robust resource allocation using edge computing for vehicle to infrastructure (v2i) networks,” in *2019 IEEE 3rd International Conference on Fog and Edge Computing (ICFEC)*, pp. 1–6, Larnaca, Cyprus, 2019.
- [85] X. Li, Y. Dang, M. Aazam, X. Peng, T. Chen, and C. Chen, “Energy-efficient computation offloading in vehicular edge cloud computing,” *IEEE Access*, vol. 8, pp. 37632–37644, 2020.
- [86] Z. Li, C. Wang, and C. Jiang, “User association for load balancing in vehicular networks: an online reinforcement learning approach,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 8, pp. 2217–2228, 2017.
- [87] J. Zhang, H. Guo, J. Liu, and Y. Zhang, “Task offloading in vehicular edge computing networks: a load-balancing solution,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2092–2104, 2020.
- [88] M. Khayyat, I. A. Elgendy, A. Muthanna, A. S. Alshahrani, S. Alharbi, and A. Koucheryavy, “Advanced deep learning-based computational offloading for multilevel vehicular edge-cloud computing networks,” *IEEE Access*, vol. 8, pp. 137052–137062, 2020.
- [89] W. Zhan, C. Luo, J. Wang et al., “Deep-reinforcement-learning-based offloading scheduling for vehicular edge computing,” *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5449–5465, 2020.
- [90] Y. Zhou, F. R. Yu, J. Chen, and Y. Kuo, “Resource allocation for information-centric virtualized heterogeneous networks with in-network caching and mobile edge computing,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 11339–11351, 2017.
- [91] C. Yang, Y. Liu, X. Chen, W. Zhong, and S. Xie, “Efficient mobility-aware task offloading for vehicular edge computing networks,” *IEEE Access*, vol. 7, pp. 26652–26664, 2019.
- [92] H. Guo, J. Liu, J. Ren, and Y. Zhang, “Intelligent task offloading in vehicular edge computing networks,” *IEEE Wireless Communications*, vol. 27, no. 4, pp. 126–132, 2020.
- [93] M. Li, J. Gao, L. Zhao, and X. Shen, “Deep reinforcement learning for collaborative edge computing in vehicular networks,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1122–1135, 2020.
- [94] H. Peng and X. Shen, “Deep reinforcement learning based resource management for multi-access edge computing in vehicular networks,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2416–2428, 2020.
- [95] H. Wang, H. Ke, G. Liu, and W. Sun, “Computation migration and resource allocation in heterogeneous vehicular networks: a deep reinforcement learning approach,” *IEEE Access*, vol. 8, pp. 171140–171153, 2020.
- [96] H. Peng and X. Shen, “Multi-agent reinforcement learning based resource management in mec- and uav-assisted vehicular networks,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 131–141, 2021.
- [97] H. Peng and X. S. Shen, “Ddpq-based resource management for mec/uav-assisted vehicular networks,” in *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pp. 1–6, Victoria, BC, Canada, 2020.
- [98] H. Wang, H. Ke, and W. Sun, “Unmanned-aerial-vehicle-assisted computation offloading for mobile edge computing based on deep reinforcement learning,” *IEEE Access*, vol. 8, pp. 180784–180798, 2020.
- [99] K. Zheng, H. Meng, P. Chatzimisios, L. Lei, and X. Shen, “An smdp-based resource allocation in vehicular cloud computing systems,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7920–7928, 2015.

- [100] M. T. Nguyen, L. H. Truong, T. T. Tran, and C. F. Chien, "Artificial intelligence based data processing algorithm for video surveillance to empower industry 3.5," *Computers & Industrial Engineering*, vol. 148, p. 106671, 2020.
- [101] M. T. Nguyen, L. H. Truong, and T. T. H. Le, "Video surveillance processing algorithms utilizing artificial intelligent (AI) for unmanned autonomous vehicles (UAVs)," *MethodsX*, vol. 8, 2021.
- [102] H. T. Do, H. T. Hua, H. T. Nguyen et al., "Formation control algorithms for multiple-UAVs: a comprehensive survey," *EAI Transactions on Industrial Networks and Intelligent Systems*, vol. 8, no. 27, 2021.
- [103] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "UAV relay in VANETs against smart jamming with reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4087–4097, 2018.
- [104] L. Zhang, Z. Zhao, Q. Wu, H. Zhao, H. Xu, and X. Wu, "Energy-aware dynamic resource allocation in UAV assisted mobile edge computing over social internet of vehicles," *IEEE Access*, vol. 6, pp. 56700–56715, 2018.
- [105] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6g communications: challenges, methods, and future directions," *China Communications*, vol. 17, no. 9, pp. 105–118, 2020.
- [106] Y. Cui, Y. Liang, and R. Wang, "Resource allocation algorithm with multi-platform intelligent offloading in d2d-enabled vehicular networks," *IEEE Access*, vol. 7, pp. 21246–21253, 2019.
- [107] Q. Pham, F. Fang, V. N. Ha et al., "A survey of multi-access edge computing in 5g and beyond: fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.
- [108] H. Peng, Q. Ye, and X. S. Shen, "Sdn-based resource management for autonomous vehicular networks: a multi-access edge computing approach," *IEEE Wireless Communications*, vol. 26, no. 4, pp. 156–162, 2019.