

A Survey on Artificial-Intelligence-Based Internet of Vehicles Utilizing Unmanned Aerial Vehicles

Syed Ammad Ali Shah , Xavier Fernando * and Rasha Kashef *

Department of Electrical, Computer and Biomedical Engineering, Toronto Metropolitan University, Toronto, ON M5B 2K3, Canada; s10shah@torontomu.ca

* Correspondence: fernando@torontomu.ca (X.F.); rkashef@torontomu.ca (R.K.)

Abstract: As Autonomous Vehicles continue to advance and Intelligent Transportation Systems are implemented globally, vehicular ad hoc networks (VANETs) are increasingly becoming a part of the Internet, creating the Internet of Vehicles (IoV). In an IoV framework, vehicles communicate with each other, roadside units (RSUs), and the surrounding infrastructure, leveraging edge, fog, and cloud computing for diverse tasks. These networks must support dynamic vehicular mobility and meet strict Quality of Service (QoS) requirements, such as ultra-low latency and high throughput. Terrestrial wireless networks often fail to satisfy these needs, which has led to the integration of Unmanned Aerial Vehicles (UAVs) into IoV systems. UAV transceivers provide superior line-of-sight (LOS) connections with vehicles, offering better connectivity than ground-based RSUs and serving as mobile RSUs (mRSUs). UAVs improve IoV performance in several ways, but traditional optimization methods are inadequate for dynamic vehicular environments. As a result, recent studies have been incorporating Artificial Intelligence (AI) and Machine Learning (ML) algorithms into UAV-assisted IoV systems to enhance network performance, particularly in complex areas like resource allocation, routing, and mobility management. This survey paper reviews the latest AI/ML research in UAV-IoV networks, with a focus on resource and trajectory management and routing. It analyzes different AI techniques, their training features, and architectures from various studies; addresses the limitations of AI methods, including the demand for computational resources, availability of real-world data, and the complexity of AI models in UAV-IoV contexts; and considers future research directions in UAV-IoV.

Keywords: Internet of Vehicles; unmanned aerial vehicles; Machine Learning; Artificial Intelligence; resource management; routing; task offloading; trajectory; survey



Citation: Ali Shah, S.A.; Fernando, X.; Kashef, R. A Survey on Artificial-Intelligence-Based Internet of Vehicles Utilizing Unmanned Aerial Vehicles. *Drones* **2024**, *8*, 353. <https://doi.org/10.3390/drones8080353>

Academic Editor: Pablo Rodríguez-González

Received: 14 June 2024

Revised: 17 July 2024

Accepted: 23 July 2024

Published: 29 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The recent surge in vehicular communication demands has given rise to the concept Internet of Vehicles (IoV). As a subset of the Internet of Things (IoT), the IoV consists of mobile vehicles outfitted with sensors, processors, and software, enabling them to communicate via the Internet or other networks [1,2]. The IoV is a decentralized system that ensures the security and privacy of both vehicular and user data, integrating various technologies to provide reliable communication tools [3,4].

Vehicle-to-everything (V2X) communication within IoV facilitates data sharing from vehicles to infrastructure (V2I), among vehicles (V2V), with pedestrians (V2P), roadside units (V2R), and unmanned aerial vehicles (V2U). This integration is vital for the intelligent management of vehicular and network data traffic, promoting safer roads and improved vehicular energy efficiency. However, V2X communication has its challenges. The high mobility and diverse densities of vehicles in the IoV necessitate continuous communication links for reliable data exchange. Fixed infrastructures, such as RSUs and BSs, often cannot provide sufficient communication and computational services, resulting in reduced QoS.

Incorporating Unmanned Aerial Vehicles (UAVs) into the IoV can significantly enhance the communication infrastructure by providing better LOS connectivity. This integration supports load balancing, mobility management, routing solutions, and cost-effective communication. UAVs, capable of autonomous operation and equipped with sensors, computing units, cameras, GPS, and wireless transceivers, can autonomously navigate pre-determined flight paths, interact with their environment, and dynamically alter their routes during flight when needed, making them a valuable addition to the IoV [5]. Specifically, UAVs can address the constraints of fixed roadside units (RSUs) by altering their speed and position dynamically, enabling them to collect and relay data across different regions [6].

The remainder of the paper is structured as follows: Section 2 summarizes the existing survey on the IoV, UAV, and UAV assisted IoV networks. Section 3 provides a foundational discussion on the principles of IoV and UAV networks with a focus on vehicular communication technologies and UAV transceiver components with UAV communication architecture. Section 4 explores AI/ML-based resource management in IoVs (Section 4.1), UAVs (Section 4.2), and UAV-assisted IoV networks (Section 4.3). Furthermore, in this section, we divide the resources in UAV-assisted IoV network into different categories, namely deployment, task offloading, trajectory, resource allocation, spectrum sharing, clustering, and energy optimization and review the research in all the categories. In Section 5, we firstly introduce the types of routing (Section 5.1) in IoV and UAV networks, namely position-based, topology-based and AI-based routing. After this, the research in the area of AI-enabled routing protocols in IoV (Section 5.2), UAV (Section 5.3), and UAV-assisted IoV networks (Section 5.4) is reviewed and critically discussed. Section 6 outlines the challenges, open issues, and prospective future research in ML/AI-based UAV-IoV networks, and finally, Section 7 concludes the paper.

2. Related Work and Survey Contribution

In the last decade, AI has been integrated into vehicular networks as a potent solution for diverse communication and traffic challenges. Coupled with V2X technology, AI enables sophisticated vehicular applications such as traffic management, Autonomous Vehicle navigation, and data management. Machine Learning (ML) and Deep Learning (DL), as prominent branches of AI, are heavily employed in the IoV to tackle complex problems by leveraging the abundant data available [7].

Most deep learning models require extensive historical data that include a variety of traffic features for training [8]. However, in vehicular communications, this historical data, which encompasses routing, channel conditions, vehicle mobility, and resources, is often not available, making supervised DL methods impractical. As a result, Reinforcement Learning (RL) has become a powerful alternative in the IoV domain, allowing vehicles to independently make decisions for various networking tasks [9–11]. In RL, agents, usually vehicles, gather data about their dynamic environment and make informed decisions to achieve goals such as resource and mobility management, routing options, and traffic forecasting.

The extensive research on AI/ML applications in wireless networks and Vehicular Ad hoc Networks (VANET) communications is thoroughly documented in scholarly articles. Liang et al. [12] investigate the application of AI/ML in analyzing mobility and traffic patterns in dynamic vehicular networks, proposing methods to improve network performance in security, handover, resource management, and congestion control. The authors in [13] categorize vehicular research related to transportation and networks, outlining vehicular network scenarios that utilize AI/ML for data offloading, mobile edge computing (MEC), network security, and transportation elements such as platooning, autonomous navigation, and safety. Furthermore, in [14], authors present a detailed review of ML techniques in vehicular networks, focusing on resource and network traffic management and reliability. This work was further extended by [15] to encompass cognitive radio (CR), beamforming, routing, orthogonal frequency-division multiple access (OFDMA), and non-orthogonal multiple access (NOMA) tasks.

An overview of ML, CR, VANET, and CR-VANET architectures including open issues and future challenges is presented in [16]. Moreover, the applications of AI/ML in CR-VANET in autonomous vehicular networks and their union are also reviewed in this paper. In [17], the authors first discussed Federated Learning (FL) and its use in wireless IoT. Then, this survey paper points out and discusses the technical challenges for FL-based vehicular IoT with future research directions. The survey paper [18] critically reviews the ML and Deep Reinforcement Learning (DRL) models for MEC decision-based offloading in IoV. The main focus of the paper is on buffer and energy-aware ML-enabled Quality of Experience (QoE) optimization, and it summarizes the recent related research and methods and presents their comparison. In [19], the authors surveyed and analyzed the resource allocation scenarios. In addition to this, the design challenges for resource management in VANETs using ML are presented as well. In [20], a detailed overview of the RL and DRL techniques in IoV networks such as joint user association and beam forming, caching, data-offloading decisions, energy-efficient management of resources, and vehicular infrastructure management is presented. Then, future trends, challenges, and open issues in 6 G-based IoV are discussed.

In [21], the primary ML concepts for wireless sensor networks (WSNs) and VANETs are summarized briefly with open issues and challenges. In [22] a comprehensive survey of AI/ML techniques is presented, and then the strengths and weaknesses of these AI models for the VANET environment, including safety, traffic, infotainment applications, security, routing, resource, and mobility management are provided. In [23] authors surveyed resource allocation techniques on DSRC, Cellular-V2X (C-V2X), and heterogeneous VANET. The AI/ML techniques are reviewed with respect to their integration in VANETs and utilization in designing several resource allocation tasks related to user association, handover, and virtual resource management for V2V and V2I communications. However, AI/ML on V2X is not the main focus of the paper.

In [24], the RL-based routing schemes are classified depending on the centralized and distributed learning process. Moreover, they surveyed position-based, cluster-based, and topology-based routing protocols. The survey in [25] summarizes the vehicular network and Smart Transport Infrastructure (STI) in detail. The paper deals with FL and its application in vehicular networks. It elaborates on vehicular IoTs (VIoTs), blockchain, FL, and intelligent transportation infrastructure. Then, the FL- and blockchain-based security and privacy applications in the VANET environment are discussed in detail. The challenges arising from the integration of FL and blockchain are pointed out in the survey with an indication of future research directions. In [26], the survey presents a compilation of network-controlled functions that have been optimized through data-driven approaches in vehicular environments. The research related to the integration of AI/ML and V2X communications in areas such as handover and resource management or user association, caching, routing, beam-forming optimization, and QoS prediction are extensively reviewed. This survey classifies the training architecture into a centralized, distributed, or federated model for each ML technique. The time complexity of supervised, unsupervised, and RL models used in the literature are discussed. In [27], the authors focused on resource management and computational offloading in a 6G vehicle-to-everything (V2X) network using FL. The paper explained the taxonomy of computational offloading in vehicular networks and only cited a few papers based on AI-driven computational offloading and focused more on explaining the different scenarios and challenges related to the network, resource management, computational offloading, and security and privacy issues in highly mobile vehicular network.

In [28], the authors cover the applications of UAVs-based IoV networks. This work does not include the detailed implementation of AI/ML in UAV-based IoV and only mentions a few papers related to Software-Defined Networks (SDNs) based fog computing and AI/ML networks. However, it covers the areas such as privacy, security, congestion and network delays, and communication protocols. In [29], the authors review the Internet of Drones (IoD) and classify the IoD-UAV according to its applications in the areas of

resource allocation, aerial surveillance and security, and mobility in all the possible IoT-based fields. This survey concludes that the most used AI technique in IoD is Convolution Neural Networks (CNNs) and the most common areas of research are resource and mobility management. However, this survey completely ignores IoD-based IoV networks. In [30], the role of UAVs in different scenarios such as smart farming and air quality indexing are discussed. One section briefly discusses the implementation of UAVs in communications as base stations, relay communication, and radio and distribution units. However, this survey does not cover UAV-assisted vehicular communication and the UAV-based resource management in depth. Similarly, in [31], the authors primarily reviewed the UAV applications in 5G network, public safety, millimeter waves, and radio-based sensing. However, there is one section of the paper that reviewed the application of ML in UAV trajectory optimization and computational offloading for 5G networks and UAV-driven federated edge learning. For computational offloading, the authors do not cite any papers and only explain the application through two diagrams. The summary of the survey papers with AI/ML applications in IoV networks is provided in Table 1.

Table 1. Summary of existing surveys.

Reference	Main Research Area	Domain	AI Technique Covered
[12]	Resource management, security and congestion control	VANET	ML and RL
[13]	Mobile edge offloading, security, transportation	VANET	ML and RL
[14]	Resource allocation, security, cognitive radio	VANET	ML and DL
[15]	Spectrum allocation	CR-VANET	ML and DL
[16]	Security, traffic safety and congestion	CR-VANET	ML, DL and RL
[17]	FL-based wireless IoT applications	CR-IoT	FL
[18]	MEC decision-based offloading	VANET	ML and DRL
[19]	Resource allocation scenarios	VANET	ML and DRL
[20]	Caching, resource and infrastructure management	IoV	DRL
[21]	Wireless sensor networks	VANET	ML
[22]	Security, routing, resource and mobility management	VANET	ML and DL
[23]	Resource allocation techniques	C-V2X	ML
[24]	Position, cluster and topology-based routing algorithms	VANET	RL and DRL
[25]	FL based security and privacy applications	VANET	FL
[26]	Handover, caching and resource management, routing	V2X Communication	ML, DL, DRL, FL
[27]	Resource management	V2X Communications	FL
[28]	Privacy, security, congestion and network delays in fog computing	UAV-IoV	None
[29]	Resource, mobility and security management and object detection	IoD	ML, DL, DRL
[30]	UAV-based resource and network management	UAV	ML and RL
[31]	UAV applications in 5G network, flying ad hoc networks and satellite, computational offloading and UAV trajectory optimization	UAV	ML and FL

In our review of existing literature on UAV-based IoV, we noted the lack of exploration into the use of ML for resource management and routing within UAV or IoD-based IoV systems. To date, no comprehensive survey has been published on this topic. Our paper provides an in-depth analysis of the integration of Autonomous Vehicles with UAVs and the deployment of AI/ML in UAV-based IoV for the allocation of physical and computational resources, as well as for routing algorithms. We discuss current AI/ML solutions in UAV-IoV, identify challenges and issues, and propose directions for future research.

3. Overview of IoV and UAV Networks

A UAV-based IoV network offers rapid data transmission services through the integration of diverse networks across various computing and communication layers [32,33]. The benefits of UAV-assisted IoV networks have accelerated their implementation in real-world scenarios. UAV-based IoV networks utilize cellular networks and communication

protocols to maintain continuous connectivity, fulfilling QoS requirements [14]. The architecture of a UAV-assisted IoV network is depicted in Figure 1, illustrating UAVs in an urban setting, aiding vehicles and infrastructure with communication services. Additionally, UAVs communicate with each other. This section presents an in-depth overview of IoV and VANET communication technologies, along with UAV components and network architecture. It also outlines the classifications and benefits of these technologies.

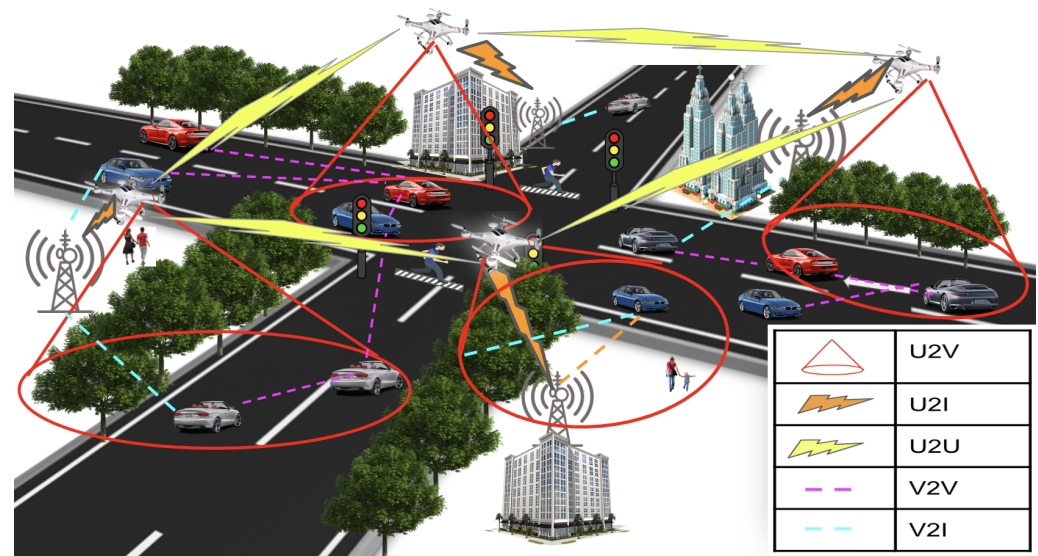


Figure 1. A UAV-assisted Internet of Vehicles (IoV) scenario.

3.1. Vehicular Communication Technologies

The IoV enables real-time communication between vehicles and infrastructure (V2I), vehicles (V2V), pedestrians (V2P), roadside units (V2R), and UAVs (V2U). Unlike VANET, which lacks Internet access and relies on Dedicated Short-Range Communication (DSRC) [34] for vehicular interactions, IoV offers a broader connectivity range. DSRC, based on IEEE 802.11p [35], was standardized in 2012 when the FCC allocated 75 MHz of bandwidth in the 5.85–5.925 GHz frequency range. It achieves latency as low as 100 ms. However, DSRC's reliance on the CSMA/CA MAC protocol limits its ability to meet the latency requirements of future vehicular communication applications with 6G and beyond, leading to potential unbounded latency and reliability issues [36].

Visible Light Communication (VLC) is an emerging technology with the potential to address the issue of spectrum scarcity [37]. Its spectrum spans from 430 to 790 THz [38]. Unlike DSRC, VLC is not affected by electromagnetic interference, offers low latency, and is less susceptible to security attacks compared to radio-based systems [39]. However, VLC does require an LOS for satisfactory performance and can be impaired by ambient light [40].

Millimeter-wave (mmWave) technology can deliver over 1 Gbps for vehicle-to-vehicle (V2V) communication and shows great promise [41,42]. Recently, mmWave-based Giga-V2V (GiV2V) has received significant attention in VANET communications. mmWave is well suited for applications that demand rich data and high definition, such as those using cameras and LiDAR sensors [43]. However, mmWave also presents challenges: it has a limited range, experiences high penetration loss, requires a line of sight, has poor diffraction capabilities, and is generally more expensive.

However, the aforementioned standards are unable to ensure steady network connectivity for highly mobile vehicles. For instance, long-term evolution (LTE) or device-to-device (D2D) and DSRC can only provide up to 100 Mbps or 3–27 Mbps, respectively [44].

In 2017, the 3rd-Generation Partnership Project (3GPP) introduced Cellular-V2X (C-V2X), which leverages the capabilities of 4G, 5G, and the anticipated 6G cellular networks [45]. C-V2X enhances safety by delivering superior system performance, extended communication range, and robust security. Within C-V2X, the PC5 [46] interface establishes

a direct communication channel between vehicles, ensuring continuous connectivity even if the link to the cellular base station is lost. Another logical interface in C-V2X is the Uu interface [47]. The IoV-6G is poised to be a revolutionary technology, addressing the limitations of 5G and meeting stringent key performance indicators (KPIs) [48].

3.2. The Unmanned Aerial Vehicle (UAV) Transceiver

The deployment of UAVs can significantly improve coverage. They act as relays, capable of transferring data from vehicles to base stations (BSs) and vice versa, enhancing the capacity of current IoV systems. UAVs can also function autonomously as BS, transmitting signals to users and increasing the system's overall capacity. Furthermore, UAVs enable better line-of-sight (LOS) communication by establishing direct aerial links with vehicles. Cellular communication requires both a BS antenna and a central switching center, reflecting a hierarchical structure. In contrast, UAV networks can be deployed on demand without such infrastructure. UAVs provide services by hovering at a specific point, ensuring vehicle coverage regardless of the scenario. This is depicted in Figure 1, where five UAV-assisted IoV networks offer coverage above buildings, pedestrians, and vehicles. Therefore, UAV-assisted IoV networks are particularly useful in densely populated areas, like sports stadiums and festivals, with many mobile users. This section discusses the fundamental components and communication architecture of UAVs.

3.2.1. Components of a UAV

An Unmanned Aerial System (UAS) comprises a UAV and a remote control system for its operation [49]. In an IoV setting, a UAV typically includes a single-board computer with a CPU, memory, and various sensors. These sensors are crucial for environmental perception, encompassing the GPS, accelerometers, cameras, and gyroscopes for navigation [50]. Additionally, a battery supplies power, a transceiver facilitates data exchange between UAVs and the Ground Control Station (GCS) [51], a flight controller manages takeoff and landing [52], and an inertial measurement unit (IMU) regulates the UAV's altitude [53]. It also has UAV flight status indicator devices [54]. In the UAV-based IoV, a UAV and a vehicle can communicate in real time typically using LoS data links. This communication capability allows UAVs to function as mobile base stations (mBSs), facilitating the transmission and reception of data between vehicles, fixed infrastructure, and other UAVs [6,55].

3.2.2. UAV Communication Architecture

UAV communication architecture is primarily categorized into centralized and decentralized types, as shown in Figure 2. In centralized communication, as depicted in Figure 3, the UAV interacts with a central controller. There are three varieties of centralized communications: firstly, UAV-GCS, where the UAV retrieves data from the GCS via a communication link, which may not be reliable in adverse weather conditions, and secondly, UAV-satellite or UAV-High Altitude Platform (HAP) communication, which is utilized for long distances between the UAV and GCS. Lastly, UAV-cellular communication, which employs cellular Base Stations to enable routing technology among nodes [32,56].

In decentralized communication architecture, as depicted in Figure 4, UAVs establish direct or indirect links with the GCS. Gateway UAVs serve as relays, transferring data between the GCS and other UAVs within the network. In ad hoc networks, UAVs communicate wirelessly with one another, independent of the GCS, as referenced in [33,57]. There are three types of decentralized UAV communication networks: simple UAV, multi-group, and multi-layer ad hoc networks. In a UAV ad hoc network, the backbone UAV links to the GCS using high power for long-range communication. This backbone UAV then acts as a gateway, connecting with other UAVs over short ranges using low power. In a multi-group UAV ad hoc network, each UAV group operates as a Flying Ad hoc Network (FANET), with one UAV designated as the backbone to communicate with the GCS. Finally, in multi-layer UAV ad hoc networks, the lower layer facilitates intra-group communication, while the upper layer consists of backbone UAVs that ensure communication between

UAVs and the GCS. The formation of multiple links in UAV-based communication aids in covering extensive communication areas, as discussed in [56].

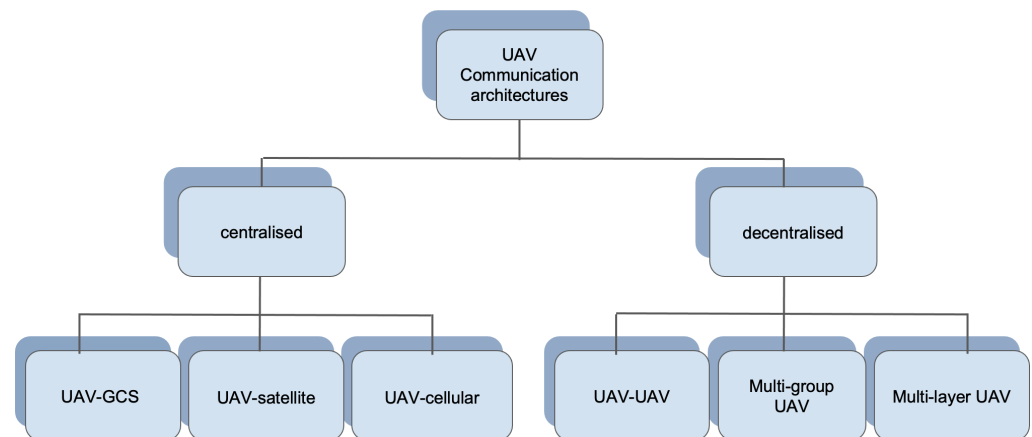


Figure 2. UAV communication architecture.

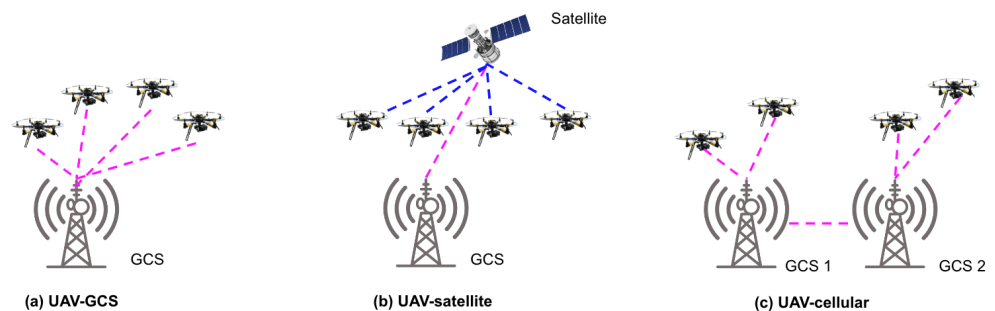


Figure 3. UAV centralized communication.

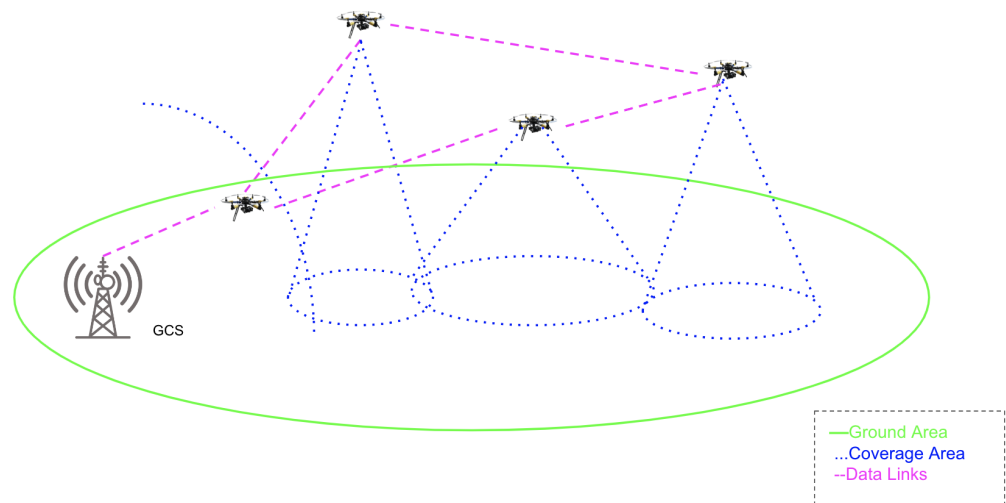


Figure 4. UAV decentralized communication.

4. AI-Based Resource Management

Addressing challenges such as high user volume, meeting stringent QoS, enhanced coverage, and cost reduction for end users requires efficiently managing various resources in UAV-assisted networks. Effective resource management is vital in overcoming challenges associated with resource scarcity. Section 4 discusses the research contributions in resource management in (1) IoV, (2) UAV, and (3) the UAV-based IoV networks utilizing AI.

The research in this area is focused on optimizing the communication resources based on the applications and services. For this reason, we divided the resource management problem into different categories based on IoV/UAV deployment, task offloading, UAV/IoV trajectory optimization, resource allocation, spectrum sharing, UAV/IoV clustering, and energy optimization, as shown in Figure 5.

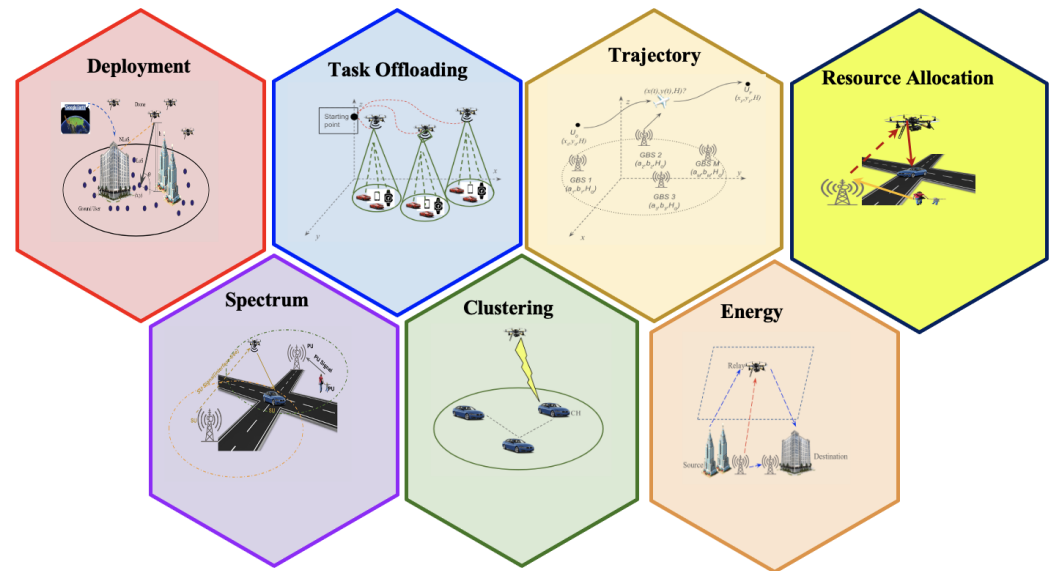


Figure 5. Categorization of resources for management in UAV-assisted IoV networks.

4.1. AI for Resource Management in Internet of Vehicles

Authors have considered various objectives for AI-based resource allocation, including load balancing, improved QoS or QoE, and the minimization of energy and latency. This section is dedicated to reviewing the contributions and research advancements in AI-based resource allocation within the IoV networks.

4.1.1. AI-Based Vehicular Clustering

In VANETs, the clustering of vehicles is performed for group nodes with similar attributes, based on predefined parameters or proximity, which facilitates the organized management of network parameters. Clustering in VANETs provides several benefits, such as resolving hidden node problems, creating manageable groups based on proximity, and efficient bandwidth utilization through frequency reuse. Typically, vehicular clustering involves designating one vehicle as the cluster head, while gateway nodes (GWs), within the transmission range of multiple cluster heads (CHs), help distribute the load. Clustering allows VANETs to leverage both wireless and wired infrastructure features effectively. Extensive research on communication protocols and strategies for ad hoc networks has identified clustering as a beneficial approach. Given the variable speeds and numbers of vehicles on the road at any time, developing a reliable mechanism for vehicle clustering is essential to evenly distribute the load on roadside units (RSUs). Supervised and unsupervised Machine Learning techniques have shown great promise in efficiently managing vehicular clusters.

In [58], the main objective is to maximize the information capacity of the VANETs by maximizing the information capacities between the head vehicle and RSU and among vehicles. Cluster-Enabled Cooperative Scheduling based on RL (CCSRL) is introduced to schedule vehicles and manage communication resources to maximize information capacity. The CCSRL primarily considers factors such as distance metrics, vehicle stability, bandwidth efficiency, velocity, density, and channel conditions to arrange the vehicles in different clusters as well as in different classes. Auxiliary vehicles are selected by considering factors

such as the vehicle's speed deviation, alignment with the direction of the CH vehicle, and the quality of the channel condition. Initially, the RSU selects the cluster head vehicle, and afterward, the new CH is selected by the previous CH. The batch size and the number of vehicles in a cluster are kept small in this research as the larger batch size prevents the RL algorithm from achieving global optimization, and as the number of vehicles increases in a cluster, the number of motion states increases and the head vehicle takes a long time to make a final decision. As a result, the convergence time of the CCSRL increases as the batch size increases. The transmission delay, throughput, and packet delivery ratio are the metrics used to evaluate the algorithm's performance.

In [59], to mitigate the effects of unreliable V2V links, each V2V pair determines whether to use the V2V mode or the V2I mode based on actual link qualities. A combined problem of selecting transmission modes, allocating radio resources, and controlling power for cellular V2X communications is defined to maximize the total capacity of V2I. A two-timescale federated DRL-based algorithm is further developed to help obtain robust models, wherein, a graph-based vehicle clustering is performed to cluster nearby vehicles on a large timescale, while vehicles in the same cluster cooperate to train the robust global DRL model through FL on a small timescale. The performance of the proposed algorithm is better than the other DRL-based decentralized learning schemes without transfer learning. However, when compared with the centralized algorithm, the proposed model does not achieve a better data rate when the number of V2V pairs increases and the outage threshold increases. Moreover, the convergence reward of the proposed model increases slowly and remains below the centralized algorithm.

In [60], the authors employ clustering by incorporating link reliability status, k-connectivity, and relative velocity factor into a fuzzy logic scheme. Each vehicle calculates a leadership value for itself and its one-hop neighbors by exchanging messages. They also employ an improved Q-learning (IQL) approach for selecting the gateway or cluster head vehicle. The proposed two-level clustering scheme demonstrates superior performance by maintaining higher throughput as vehicle speeds increase and reducing the likelihood of route changes compared to other classical Q-learning algorithms. However, the authors acknowledge the fact that as the action and state space grow, the complexity and computational cost of the proposed algorithm are expected to be increased significantly, and they do not deal with the complex vehicular scenario.

The authors in [61] use Q-learning to select the optimal next-hop grid. The authors used grid-based routing, which divides a geographic area into small grids, allowing Q-learning to determine the optimal sequence of grids from source to destination. Once the optimal grid is selected, the agents choose the best relay vehicle within that grid using a greedy or Markov prediction method. Buses are given higher priority in vehicle selection because of their fixed routes and schedules, enhancing the scheme's performance. Simulations indicate that this hierarchical routing scheme improves the delivery ratio and throughput, although it results in similar or slightly increased delay, hop count, and packet-forwarding frequency compared to other position-based routing protocols for various time slots. The authors do not provide any information about the complexity and control information overhead of the proposed model. It is an important aspect of the research problem as the proposed protocols need extra overhead for computing the Q-table compared to other comparative protocols.

In recent years, research has shifted from supervised learning to Q-learning-based RL algorithms to address clustering issues in VANET for time-sensitive applications. It is observed that in references [58–61], the focus is on employing Q-learning with an RL algorithm, utilizing a limited action and state space to maintain a small Q-table and reduce computational expense. However, V2V and V2I communications are time-critical, and large Q-tables can be impractical for time-sensitive applications. Consequently, there is a need to investigate Deep RL algorithms for VANET clustering to manage the increasing complexity as the action and state space expand. Deep Q-Networks (DQN) use neural networks to replace the Q-table, taking the state as input and predicting Q values based on historical

data. The research work in the area of vehicular clustering for resource management is summarized in Table 2 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 2. AI/ML solution for IoV-based clustering.

Reference	Objective	Algorithm	Metrics
[58]	information capacity of VANET	CCSRL	transmission delay, throughput and packet delivery ratio
[59]	maximize the total capacity of V2I	two timescale federated DRL-based algorithm	sum capacity of V2I and satisfied rate of V2V pairs
[60]	throughput maximization	fuzzy logic based improved Q-Learning	route request message, throughput,
[61]	next hop grid	Q-learning	delivery ratio and throughput

4.1.2. AI-Based Vehicular Spectrum Sharing

Spectrum sharing involves managing the distribution of spectrum among Vehicular Cognitive Radio (VCR) users while ensuring QoS. It can be categorized based on spectrum utilization into unlicensed and licensed types. In unlicensed spectrum sharing, all users have equal priority, while in licensed spectrum sharing, primary users (PUs) are prioritized over secondary users (SUs). SUs can access both types of spectrum sharing only when PUs are not utilizing the spectrum. Additionally, spectrum sharing can be classified as centralized or distributed. In centralized spectrum sharing, a central node controls spectrum allocation and access, while in distributed spectrum sharing, each node independently manages spectrum access. Cooperative and non-cooperative approaches are also employed in spectrum sharing within VCR networks.

In [62], the authors use the multi-agent RL (MARL) to develop a distributed spectrum sharing and power allocation algorithm to enhance the performance of both V2V and V2I links together. In the RL environment, the multiple V2V links try to access the V2I spectrum. The V2V links act as agents and refine the spectrum allocation and power control—strategies based on the individual environmental observations. Instead of considering the continuous values for power control, this paper considers only four levels of power control. This eases the learning by reducing the dimensions of the action space. In the training stage, the proposed model is centralized, and in the implementation stage, it is decentralized. The proposed MARL and single-agent RL (SARL) algorithms are used for comparison purposes. The proposed model considerably improves the overall system-level performance.

In [63], the authors used the same MARL-based approach and four-level power con dimensionality reduction for the action space in NOMA communication using the MARL algorithm. So, in addressing spectrum allocation issues in V2X communications, the objectives are to enhance the overall throughput of V2I links while increasing the probability of the success of V2V channels within a specified time constraint, T . However, the reward function in this study is not defined, and no reward penalty is provided. Moreover, the complexity of action and state space are not elaborated. The convergence of the reward function is not provided, so, it is difficult to draw conclusions about the performance of the proposed algorithm.

In [64], the authors address three issues of reliable Cooperative Spectrum Sensing (CSS), channel indexing for selective Spectrum Sensing (SS), and optimal channel allocation to CR SU in a single framework for CR-VANETs. For CSS, local SS decisions with critical attributes such as the geographical position of the sensing signal acquisition and timestamp, utilizing the DRL technique to obtain a global CSS session, are introduced. All the vehicles (static and mobile) and UAVS are considered as SUs. Selective channel-based spectrum sensing is employed to reduce the sensing overload on CR users. A time series analysis is used with a deep learning-based Long Short Term Memory (LSTM) model to index PU channels for selective SS. Finally, for channel allocation in CR-VANETs, the complex

environment is modelled as a Partially Observable Markov Decision Process (POMDP) framework and solved using a value iteration-based algorithm. To reduce the dimensionality problem associated with the DRL algorithm, the approximation method is used to reduce the size of the action and state space in the proposed algorithm. The reward function formulated in this research is highly unstable and only stabilizes for a few episodes and again drops and starts fluctuating. Moreover, the probability of PU detection drops as the speed of vehicles increases.

In [65], the proposed resource management mechanism achieves intelligent and dynamic control of the entire VANET. The BS of each cell acts as the DRL agent. The environment encompasses the entire vehicular communication network, including the BS, IRS-aided channel, and the vehicles. The objective of the DRL-based scheme is to jointly optimize the transmission power vector of head vehicles, the Intelligent Reconfigurable Surface (IRS) reflection phase shift, and the BS detection matrix to maximize network energy efficiency under given latency constraints. The CSI and the status of the VANET are collected and sent to the DRL agent, which then takes action and receives the corresponding reward from the environment. Since the state and action variables of the DRL-based resource control and allocation scheme are continuous, the Deep Deterministic Policy Gradient (DDPG) algorithm is employed to solve the optimization model. The proposed model performs better than the baseline models in terms of system energy efficiency. However, the complexity of the model increases with the increase in the number of neurons in actor and critic networks, and it needs to be multiplied by the number of episodes and the number of time slots used in each episode. The comparison of the proposed algorithms' convergence is not provided to gauge the efficiency of the proposed scheme.

In [66], a resource allocation problem is solved for V2X communications aimed at maximizing the sum rate of V2I communications while ensuring the latency and reliability of V2V communications. This is achieved through a joint consideration of both frequency spectrum allocation and transmission power control. The authors formulate the resource allocation problem as a decentralized Discrete-time and Finite-state Markov Decision Process (DFMDP), where V2V links are the agents, the local channel information such as V2V-link interference channels from other V2V links are the state, the action space is the power allocation and spectrum multiplexing factor of V2V and V2I links, and finally the reward function is based on the sum-rate of V2I communications and the delivery probability of V2V communications, as the aim of the research is to select the proper spectrum bands and transmission powers that optimize the different QoS requirements of both V2I and V2V communications. To handle the continuous action space, the authors implemented a Deep Neural Network (DNN)-based DDPG framework, and higher efficiency is achieved compared to the random resource allocation scheme for the sum-rate of V2X communications and the delivery probability of V2V communications.

In [67], for V2V communication, the authors proposed the RL-based decentralized resource allocation mechanism. It is applied to both unicast and broadcast scenarios. V2V links are agents and, based on the minimum interference, select their spectrum and power transmitted for V2I and V2V links. The V2I capacity maximization and V2V latency are chosen to show the performance of the proposed algorithm. As the number of vehicles increases, the interference grows, which lowers the V2V link capacity and makes it hard to guarantee the latency. In the proposed solution, the transmitted power is divided into three levels, and the agents select them based on their state information. The DRL is able to autonomously determine how to adjust power levels based on the remaining time and intelligently allocates resources based on local observations, resulting in significant improvements in the V2V success rate and V2I capacity compared to conventional methods. The authors modify the action state (action taken by each agent), and the agents update their actions asynchronously, with only one or a small subset of V2V links updating their actions in each time slot. This approach allows agents to observe environmental changes caused by the actions of other agents.

In [68], the authors proposed a spectrum resource management multi-hop broadcast protocol named the Global Optimization algorithm based on Experience Accumulation (GOEA) to facilitate the coordination among vehicles in channel selection, aiming to mitigate packet loss resulting from channel collisions. Moreover, the dynamic spectrum access model is proposed based on the RL and Recurrent Neural Network (RNN+ DQN) algorithm. It is noted that as the number of vehicular users increases compared to the number of channels, the proposed RNN+DQN models' performance deteriorates significantly. Both the proposed DRL model and GOEA only perform better if the number of users is low; otherwise, both perform badly as the vehicular density increases.

In [69] the authors further extend their work [70] and maximize the spectrum efficiency based on the mobility-aware, priority-based channel allocation method using DRL, where channels are allocated to vehicles based on their Service Mobility Factor (SMF) and priority. LSTM networks are employed to capture the temporal variation in service requests due to user mobility, which is then integrated with DRL. The bandwidth allocation policy is optimized using the proposed algorithm. The reward is calculated based on the user's SMF, transmission cost, and used bandwidth. Additionally, the performance of LSTM+DQN and LSTM+A2C correlates with reward function convergence and spectral efficiency. Both models show superior performance for reward convergence. This study used the real-time large vehicular speed dataset, and the proposed model handles the big data efficiency. However, the loss incurred by the proposed model keeps fluctuating because the environment keeps changing.

Research on spectrum sharing extensively employs RL algorithms and their variants. The primary goal is to maximize throughput and spectral efficiency while minimizing latency for vehicular users. DQN-based algorithms are utilized to manage the continuous action and state spaces. However, as vehicle numbers and the complexity of these spaces grow, RL algorithm performance tends to decline [69]. To address this, clustering approaches have been implemented to stabilize the system and lower access latency by decreasing direct connections to the cellular network [65]. Moreover, distributed approaches are favored in the literature because they reduce message overhead among vehicles compared to centralized methods. The research work in the area of spectrum sharing is summarized in Table 3 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 3. AI/ML solutions for IoV-based spectrum sharing.

Reference	Objective	Algorithm	Metrics
[62]	Spectrum allocation and power control	MARL	V2V transmission rate and payload
[63]	Bandwidth allocation scheme based on the game theory	MARL and DDPG	Throughput
[64]	Task offloading and security assurance	LSTM+POMDP	Probability of PU detection and PU collision with CR-VANETratio
[65]	Capacity maximization using vehicle platooning	DDPG	System energy efficiency
[66]	latency	DFMDP	sum rate of V2X and delivery probability of V2V
[67]	Optimal spectrum selection and transmitted power	DRL	V2I capacity maximization and V2V latency
[68]	Channel selection, minimization of packet loss	GOEA (RNN+DQN)	Packet loss and collision probability
[69]	Channel allocation	LSTM+DQN and LSTM+A2C	Spectral efficiency

4.1.3. AI in Ground Trajectory Management

Vehicular ground trajectory prediction is vital for the safety of intelligent self-driving vehicles. It predicts traffic behaviour on the road and informs future maneuvers based on these predictions, including drivers' responses to sudden trajectory changes. Additionally, changes in vehicle trajectory and position impact the Field of View (FOV) in V2V communication as road blockage and vehicle density increase. Consequently, researchers are concentrating on trajectory prediction and have proposed numerous effective AI-based methods.

In [71], the authors predicted leading vehicle trajectory using the proposed method based on the joint time-series modelling approach (JTSM). The proposed model is compared with the constant Kalman filter (CKF), LSTM, and multiple LSTM (MLLSTM). The proposed model shows significant improvement in terms of root mean square error (RMSE). In [72], the authors predict vehicles' trajectory by using the LSTM algorithm. Then, the predicted value is provided to the QL algorithm to figure out the optimal resource allocation policy for the nodes. The real-world vehicle trajectory data used in this research were provided by Didi Chuxing, a ride-sharing company. The ultimate goal is to enhance the QoS for non-safety-related services in MEC-based vehicular networks, and the proposed model outperforms the other models.

In [73], the paper employs an LSTM encoder to encode the states of the target vehicle, enabling the prediction of its maneuvers. Trajectory prediction is then achieved using the predicted maneuvers along with map information. Finally, based on interaction-related factors, traffic rules (such as red lights), and map information, nonlinear optimization methods are utilized to refine and optimize the initial future trajectory. In addition to this, with the advancement of neural networks, various RNN architectures have been extensively utilized.

In [74], the authors employed two groups of LSTM networks to predict the trajectory of a target vehicle. One group is used to model the trajectories of surrounding vehicles, while the other group focuses on modelling the interactions between these surrounding vehicles. In [75], TraPHic, a model based on the CNN-LSTM hybrid network to predict the trajectories of traffic participants, is proposed. This model inputs the state and surrounding objects of the main vehicle into CNN-LSTM networks to extract their features. These features are then combined with the LSTM decoder to predict the main vehicle's trajectory. However, this algorithm only predicts the trajectory of one object per operation. Similarly, in [76], the authors employ a CNN-LSTM framework using a "box" method to detect and eliminate outliers in vehicle trajectories to obtain valid data. These data are then processed through the convolutional and maximum pooling layers to extract interaction-aware features, which are subsequently fed into an LSTM and a fully connected layer for prediction. The model's hyper-parameters are optimized using the Grid Search (GS) algorithm.

This research primarily focuses on applying supervised learning techniques to the vehicular ground trajectory, a subject extensively studied within autonomous vehicular environments. The key areas of interest include driving-style prediction [71], driving maneuvers [73], and trajectory prediction for safe driving [74,75]. However, the impact of vehicular trajectory on resource management in ad hoc vehicular networks remains under-explored. Vehicular trajectory prediction is crucial, as the channel condition between vehicles in a highly mobile network depends on the LOS, and even minor variations in V2V communication can significantly degrade channel conditions, affecting the overall system performance. Additionally, since VANETs operate as multi-agent networks with independently moving agents, a minor positional shift of one agent can influence others. RL algorithms show promise in predicting trajectory effects on the system, but further study is needed to understand the impact of trajectory control on resource management using Machine Learning. The research work in the area of vehicular ground trajectory optimization for resource management is summarized in Table 4 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 4. AI/ML solution for IoV-based ground trajectory management.

Reference	Objective	Algorithm	Metric (s)
[71]	Prediction of leading vehicle trajectory	JTSM	RMSE
[72]	Prediction of vehicles' trajectory	LSTM	Prediction accuracy
[73]	Trajectory prediction using vehicle's maneuvers	LSTM	RMSE
[74]	Trajectory prediction	LSTM	MSE and RMSE
[75]	Traffic trajectory prediction	CNN+LSTM	RMSE
[76]	Vehicular trajectory prediction	CNN+LSTM	MAE and RMSE

4.1.4. AI in Task Offloading

Task offloading refers to the transfer of data from one device to another or the migration of the end user from one communication network to another. This paper focuses on the offloading of data or tasks between devices to distribute network resources and balance the load in highly mobile vehicular networks. Nowadays, vehicles are outfitted with sensors, cameras, transceivers, and onboard computing devices to facilitate communication with other vehicles and the surrounding infrastructure. Vehicular offloading involves transferring or migrating computations to cloud or fog nodes to augment vehicle capabilities. This offloading process allows for the remote processing of vehicular applications within the cloud or fog infrastructure. When vehicular application computations take place at the cloud level, it is known as cloud computing. Alternatively, when vehicular applications demand low latency and high computational resources, computations are performed on fog-level servers, a practice known as fog or edge computing. Task offloading can be binary or partial. In binary offloading, the entire task is either executed locally by the vehicle or transferred to the fog server for execution. With partial offloading, the vehicle performs a portion of the task locally, while the remainder is offloaded to the vehicular edge server for completion.

In [77], the authors proposed a multi-platform intelligent offloading and resource allocation algorithm to dynamically organize the computing resources. The task-offloading problem is dealt with as a multi-class classification problem where the K-Nearest Neighbor (KNN) algorithm selects the best option available out of cloud computing, mobile edge computing, or local computing platforms. The system makes decisions to compute the complete task locally or decides to offload it to the MEC or the cloud. In addition, when the task is offloaded to a desired server, RL is implemented to solve the resource allocation strategy. The state is defined as the MEC computing capacity, the actions are the offloading decision and computation resource allocation, and the reward is the minimum total cost. The proposed joint optimization is compared with full MEC and full local techniques, and it is concluded that the proposed scheme reduces the total system cost and optimizes the overall system performance. However, the proposed RL model is not compared with any other AI or conventional mathematical optimization techniques.

A study similar to [77] is conducted by the authors in [78]. They approached the task-offloading problem in the same manner by proposing two offloading layers. The first layer selects between cloud computing (CC) and MEC servers using the Random Forest (RF) model to decide between local, MEC, or CC task offloading. In the proposed DRL model, vehicles send their traveling state, location, and task information to the MEC server. The RSU/BS is responsible for collecting the MEC server status, managing spectrum and computing resources among vehicles with task offloading requests, and combining this information into an environmental state. The RSU/BS then sends the combined environment state to the agent. The agent receives feedback on the optimal policy for resource allocation decisions for each vehicle to maximize the total accumulated reward. The T-Drive trajectory dataset, from the Microsoft website, is used for the model training and testing. This study is limited in scope, and the reward function is not well formulated, which significantly harms the convergence of the reward function.

In [79], the authors divide regions into different vehicular fog cloud (VFC) systems, and each VFC consists of moving vehicles, a remote cloud, one or more VFs, and a VF resource manager (VFRM). The VF has restricted resources, and VFRM controls the assignment of the resources to VF to fulfill service latency requirements. The authors deal with the offloading problem as partial offloading. To implement the proposed strategy, the proximal policy optimization-based RL algorithm is used to handle the continuous action space instead of Q-learning as it is not suitable for this purpose. In terms of computational time, the proposed proximal policy optimization RL (PPO-RL) model does not perform better than the simple PPO algorithm as the proposed RL model needs the computations for the heuristic model used with the proposed model.

In [80], to tackle the problem of the scarcity of computational resources, a selection criterion is proposed to select volunteers' vehicles capable of executing the computationally intensive task. For the volunteer vehicle identification or the task-offloading decision, the authors used various state-of-the-art ML-based regression techniques, including LR, SVR, KNN, DT, RF, GB, XGBoosting, AdaBoost, and ridge regression. For the training and testing of the models, a vehicular onboard unit computing capability dataset is collected. It contains three different datasets. All three datasets have seven features but a different number of samples or sizes. The results for the task execution time and delay are also made to conform with the simulation environment developed using the NS3 simulator. One drawback of the proposed scheme is that it is not delay-tolerant, whereas the computing and transmission delays in task offloading are very critical to consider.

In [81], a model named ARTNet is proposed to make an AI-enabled V2X framework for maximizing resource utilization at the fog layer and minimizing the average end-to-end delay of time-critical IoV applications in a distributed fashion. The software-defined network (SDN) controller selects the secondary agents who, in the case of SDN failure, support the underlying architecture. Moreover, the ARTNet, implemented with the secondary agents, takes the data offloading decision to minimize the end-to-end delay based on the reward function. The energy consumption, average latency, average overload probability, and energy shortfall are considered as evaluation metrics. The ARTNet model achieves success through lower latency, reduced energy consumption, and minimized energy shortfall by intelligently distributing tasks at the fog layer using resource pooling. Additionally, ARTNet assigns tasks to fog nodes with fewer tasks and optimizes the performance. However, the proposed model is a simple Q-learning-based RL model whose reward convergence analysis is not provided. The authors could have used other heuristic algorithms in the RL model instead of Q-learning for comparison purposes, or they could have compared it with other variants of RL models to present the effectiveness of the proposed model.

In [82], the authors perform queue-length resource allocation. At the controller level, the network safety flows are managed. The safety flows have a higher priority ratio based on the criticality, and the non-safety flows have less priority. The bandwidth allocation is the main fairness allocation criterion to obtain the maximum rate for different applications. The simulation environment uses mininet-wifi for multiple RSUs and vehicles to communicate in V2V and V2I scenarios. The authors implemented LSTM, CNN, and DNN and compared their results with one another. The LSTM outperforms all models in terms of accuracy. AI-supervised learning is implemented, but the study provides limited information about the data collection, the size of the data samples, and the features used to classify the flows.

In [83], the authors introduced automated slice resource control and updated the management system using two ML models. The first ML model predicts future resources at the network edges based on the user traffic streamed at each edge, classifying the traffic type to determine the specific resources required at any given physical resource location. The second ML model focuses on the resource utilization of virtual machines (VMs). It predicts future resource usage to decide on scaling specific types of virtual network functions (VNFs), ensuring service availability. The RNN model is used to automate the resource management for the IoV network and compares it with Auto-Regressive Integrated

Moving Average (ARIMA) model in terms of accuracy. The dataset used in this research is called the GWAT-13 Materna dataset with 12 attributes, available on Materna 13, an open source directory. It has three traces expanded over three three-month period with each trace having 850 VMs data on average. The prediction results of ML models are used by the Automated Slice Resource Control and Update Management System (ASR-CUMS) to decide the resource requirements and update the physical resources. The dataset used is synthetic and is used for the reliable prediction of network resources available.

In [84], the authors framed the computation offloading problem as Multi-agent Deep Reinforcement Learning (MADRL), aimed at selecting the best MEC server to execute tasks for multiple vehicles. Each vehicle's state, which includes real-time location and task information, is considered. The objective is to minimize the total task execution delay across the entire system over a given period, and the task execution delay is used as the primary performance metric. Initially, the data center trains the actor and critic networks in a centralized manner. Subsequently, vehicles make task-offloading decisions in a distributed way. The reward function is formulated as the task completion time when vehicles offload the task. However, the penalty associated with the wrong action taken by the agent is not given. The evaluation of the proposed scheme shows that MDRCO achieves superior performance compared to the NN algorithm and the AC algorithm.

In [85], the authors proposed a Lyapunov-optimization-based Multi-Agent Deep Deterministic Policy Gradient algorithm (L-MADDPG) for task offloading and resource allocation with the ultimate objective of minimizing the system energy under the queue stability and latency constraints of the vehicular network. The authors adopt a binary offloading approach to offload the task to MEC. Each vehicle keeps the local computation queue and offloading task queue. The MADDPG determines the best possible offloading policies based on the computational ratio and queue length at the edge serve. The state space includes the speed of the vehicle, the computational resources available, and the maximum available power at the vehicle. The action space includes the task-offloading policy, the local resources allocated to the task, and the size of the offloading task. The reward function is based on the amount of energy consumed in the local processing of the task and includes all the computational ratios and energy constraints. The proposed L-MADDPG is compared with other state-of-the-art RL-based algorithms. The reward function's convergence for the proposed model and simple MADDPG models are the same. For energy consumption, the proposed model outperforms the other models. However, as the number of vehicles grows, the local computation at VEC grows to 4 times the number of vehicles. The study does not address time complexity or the slow convergence of the proposed algorithm which is evident from the results presented.

Most studies have focused on binary task offloading, often neglecting partial offloading. For example, the binary offloading issue is addressed in [77,78]. In [77]: the offloading problem is tackled using a state-of-the-art RL algorithm, which struggles with computational time and the dimensions of the action and state space. To address these limitations, Ref. [78] employs DRL, which manages dimensionality and enhances computational complexity. The technique of partial offloading is utilized in [79], where Q-learning is replaced with PPO within the RL model to manage complexity. However, this approach is not benchmarked against DDPG, which is known to yield better outcomes in partial offloading scenarios with high-dimensional action spaces. Lyapunov optimization, a well-established method for task offloading, has yet to be fully explored with AI to address the intricate time and computation complexities of the task-offloading problem. In [85], the authors successfully integrate Lyapunov optimization with an RL algorithm, indicating a promising but under-explored area that warrants further research attention beyond merely the energy consumption of the system. The research work in the area of vehicular ground trajectory optimization for resource management is summarized in Table 5 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 5. AI/ML solution for IoV-based task offloading.

Reference	Objective	Algorithm	Metric (s)
[77]	computational resource allocation using task offloading to minimize the delay	KNN and RL	Total delay cost
[78]	computational resource allocation using task offloading to minimize the delay	RF and RL	Total delay cost
[79]	Fulfill latency requirements	PPO-RL	MSE, fog capacity prediction, and service satisfaction
[80]	computational resource optimization	LR, SVR, KNN, DT, RF, GB, XGBoosting, AdaBoost and ride regression	task execution time and transmission delay
[81]	minimize average end-to-end delay of time critical applications	Q-learning based RL	Latency, energy consumption and overload probability
[82]	queue-length resource allocation	LSTM, CNN and DNN	Precision, recall, F1 Score and accuracy
[83]	predicts future resource usage to scale VNF	RNN	Train and test accuracy
[84]	channel selection, minimization of packet loss	MADRL	Number of packets delivered and delay time
[85]	Minimize system energy and latency	Lyapunov based MADDPG	reward function based on energy consumption in processing the task

4.2. AI for Resource Management in UAV Networks

The mobility and LOS links offered by UAVs present them as a viable alternative to fixed base stations in wireless communication networks. Likewise, AI has garnered significant interest in this field due to its capacity to learn from data and the environment, enabling autonomous decision-making. Consequently, the research community is actively pursuing the integration of intelligence into UAV networks through various AI algorithms. This section discusses the potential applications of AI in UAV-based wireless networks, which can serve as a foundational platform for UAV-based vehicular networks.

4.2.1. AI in UAV Deployment

The placement of UAVs is critical in resource management, affecting transmit power, coverage, and the QoS of the communication system. UAVs may be deployed in various configurations, including two-dimensional, three-dimensional, single-UAV, and multi-UAV formations. In a two-dimensional setup, the UAV's altitude is fixed, whereas a three-dimensional approach takes into account all three spatial coordinates. Optimizing UAV placement has produced enhanced outcomes, which are elaborated upon in this section.

In [86], the authors combined the features of FL and MARL for UAV deployment and resource allocation in urban areas using a multi-agent collaborative environment learning (MACEL) with the main goal of enhancing the overall utility of the multi-UAV communication network through strategic adjustments in the positioning, channel allocation, and power configuration of individual UAVs. In this paper, the individual cumulative reward obtained by a single UAV with MACEL is not better than the MADQL network as each UAV in MADQL only pursues its own reward maximization and has no cooperative relationship of information sharing with others as in MACEL. Moreover, as the number of users increases, the proposed model increases the UAVs' power while adjusting the UAVs' locations to mitigate the effects of interference and energy consumption. Similarly,

as the number of UAVs is increased (for UAVs = 6), the co-channel interference increases, and although the MACEL optimizes the UAV deployment, the interference does not reach satisfactory levels. One solution to overcome the interference and network capacity optimization could be to increase the discrete number of power levels but, this will affect the complexity and computing of the network as the pace of action grows with it.

In [87], the authors find the optimal placement of each UAV-BS that minimizes energy consumption. The load prediction algorithm (LPA), which is based on two supervised ML algorithms, namely RF and Generalized Regression Neural Network (GRNN), is used to predict macro-cell congestion based on the load history generated by the mobile network. Then, the UAV-BSs Clustering and Positioning Algorithm (UCPA) is implemented to calculate the required quantity of UAV-BSs for each congested macro-cell to minimize the corresponding user congestion, alongside identifying the optimal placement of each UAV-BS within the coverage region of the congested macro-cell. The proposed model demonstrates better overall throughput, signal-to-noise ratio (SNR), and number of users supported by UAV-BS. This study comprehensively covers the UAV and non-UAV-based congestion control by setting up the simulated network with real-time data and evaluating the performance of the system under minimum throughput and SNR requirements. However, the overhead reduction during the higher user demand needs to be investigated further for the proposed intelligent system to be implemented in 5G and 6G networks, which require shorter delays and higher throughput.

In [88], the authors proposed an approach to the deployment of multiple UAVs-Reconfigurable Intelligent Surface (RISs) (RISs installed on UAVs) serving multiple downlink users. This paper jointly optimizes the active beamformers at both the macro and small-cell base stations, the phase shift matrix at each RIS, the trajectories/velocities of UAVs, and sub-carrier allocations for micro- and mmWave transmissions, with the objective of minimizing the overall transmit power of the system. The fundamental problem is non-convex and is a mixed integer programming problem, so it is decomposed into two distinct sub-problems. The first sub-problem focuses on optimizing the trajectories/velocities of UAVs, the phase shifts of RISs, and sub-carrier allocations for micro-wavelength transmissions, which is solved using the dueling-DQN learning approach by developing a distributed algorithm, while the second deals with the design of active beam-forming and sub-carrier allocation for mmWave transmissions, which is solved using the SCA method. The performance of the proposed model is compared with other baseline algorithms in terms of transmit power against the minimum data rate (increases with optimized location), the number of reflecting elements at the RISs (transmit power decreases as the number of reflecting elements at the RISs increases) and number of antennas at the MBS (the transmission power of the system efficiently decreases as the number of antennas increases). However, in this study, the number of UAVs is fixed to 2 and the effects of the large number of UAVs on the power requirements, interference, and SNR are not considered in this study.

In [89], the authors proposed a Multi-objective Joint DDPG (MJDDPG) algorithm to maximize the aggregated data collection and energy transmission within the urban monitoring network while simultaneously minimizing the energy expenditure of the UAVs and optimizing the UAV flight patterns. The results show that the data collection and amount of energy transfer by the UAVs fluctuate a lot throughout the training phase. Moreover, as the number of nodes increases, the energy consumption of UAVs in the case of the proposed model deteriorates as compared to other baseline models. In this study, the experimental results of the proposed model are not compared with any other models to prove the authenticity of the model. In addition to this, it is noted that the efficiency and validity of the designed reward function can be exploited mathematically to improve the performance of the proposed algorithm. The dimensionality effects of the large action and the state space are not considered, and the time and computational complexity of the algorithm are not discussed.

In [90], the authors consider a swarm of HAPSs for communication and aim to compare the RL and swarm intelligent (SI) algorithms. In the SI algorithm, the HAPS support a

fixed number of users, and one of the HAPSs does not support any user at all. In the RL algorithm, the number of users supported by the HAPSs changes dynamically, all HAPSs support a number of users, and the total number of users supported by all the HAPS with RLs is significantly higher than in the SI algorithm. The scope of this study is very limited and it does not cover the complicated HAPS scenarios with hybrid solutions/algorithms nor compare the results with other baseline algorithms.

In wireless communication research, UAV deployment is considered an optimization sub-problem alongside others such as user-UAV association and UAV transmit power [86], as well as energy optimizations [87]. The primary goals include maximizing QoE, sum rate, network throughput [87], the lifetime, the fairness, and the spectrum efficiency [88]. The three-dimensional deployment of UAVs poses a significant challenge in UAV-based communications and has not been thoroughly explored. Furthermore, the optimization problem of determining the 3D locations for UAV-BSs is NP-hard and lacks a deterministic polynomial-time solution. Heuristic and numerical methods have been used to approximate the optimal locations for UAV-BSs. Additionally, collision avoidance and accurate channel estimation are critical areas that must be addressed with UAV deployment in cellular networks using RL-based algorithms. Currently, AI-based solutions yield overly optimistic results in offline/simulated environments, highlighting the need for their implementation in realistic communication settings. The research work in the area of UAV deployment for resource management is summarized in Table 6 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 6. AI/ML Solution for UAV Deployment.

Reference	Objective	Algorithm	Metric (s)
[86]	overall utility enhancement of UAV communication	FL and MARL	co-channel interference and network capacity
[87]	optimal UAV deployment to minimize energy consumption	RF and GRNN	overall throughput and SNR
[88]	UAV deployment to optimize the active beamformers	Dueling-DQN	transmit power vs. minimum data rate, amount of data collected
[89]	maximize the aggregated data collection and energy transmission	MJDDPG	and UAV energy consumption
[90]	optimal HAPS deployment to support more users dynamically	RL	number of users supported.

4.2.2. AI in UAV Spectrum Management

Networks that utilize the spectrum combine aerial UAVs and terrestrial communication devices, which depend on the allocated spectrum for various tasks such as information transmission and data relaying. These networks function under three spectrum-sharing paradigms: overlay, underlay, and interweave. In overlay mode, UAVs gain access to extra bandwidth for their transmissions while supporting terrestrial transmissions. Underlay mode permits multiple nodes to concurrently share the same band while strictly managing mutual interference. In interweave mode, UAVs opportunistically transmit information when terrestrial signals are absent. The effective spectrum-sharing strategies allow both UAVs and terrestrial devices to improve their communication capabilities. UAVs can connect to terrestrial access points for high data rate and secure transmissions, and also act as aerial access points to bolster terrestrial communication.

In [91], the authors proposed a dynamic information exchange management scheme in a UAV network based on LSTM and the DQN algorithm to improve the average collision rate, throughput, and reward function based on the frame rate, sending bit rate, and total packet error rate. The performance of the proposed LSTM+DQN model is not considerably better than the DQN and Q-learning model for average packet collision rate. For the

throughput of the dynamic time slot allocation system, the proposed model converges slowly along with other comparative models and does not show better throughput maximization as compared to the other models. It is noted that the action space of each UAV agent either shares information with all other UAVs or waits to share. This makes the action space grow as all UAVs are exchanging information at the same time. In addition to this, the action space is defined as a binary operation and the best channel allocation factor is a continuous time function. Clearly, this makes the proposed model more complex and slows down the convergence, which is evident based on the results obtained in the research.

In [92], the authors proposed a DQN-based task offloading and channel allocation scheme with the objective of gathering the expected data packets while ensuring they meet the delay constraints for each packet and keep the data computational processing and time processing costs to a minimum. The volunteer vehicles need to choose the right number of sub-channels to minimize the task uploading delay, while the UAV must select the most efficient task processing model. This is a complex integer programming problem, and to solve it, the Lagrange Duality Method and DQN are deployed. The mean cost (data preparation, transmission, calculation, and downloading) is evaluated based on transmission power, the velocity of UAVs, the computing capacity, and the distance between vehicles and UAVs. The proposed DQN scheme performs better than the other Q-learning-based techniques in terms of the convergence of reward value and the computational time except for the DQN-based Double-Option Scheme, as both use neural networks to predict Q-value. Moreover, as the number of vehicles increases, the cost of the system also increases, which is the drawback of the proposed model as it is unable to serve all the vehicles simultaneously.

In [93], the authors perform joint power allocation and scheduling for a UAV swarm network. In the network, one drone is selected as a leader, and all other drones are made to be a group of drones following the leader. Every group transmits the update of its local FL model to the leader drone so it can combine all the local parameters for global parameter updates to the global model. While the drones exchange updates, the wireless transmissions are affected by many internal and external losses and interference. In order to assess the influence of wireless variables such as fading, transmission delay, and UAV antenna angle variations caused by environmental factors like wind and mechanical vibrations on FL efficiency, a comprehensive convergence analysis is conducted. Subsequently, a strategy for joint power allocation and scheduling is introduced to enhance the convergence speed of the FL. One drawback of the study is that as the variance of the angle deviation increases, the FL convergence takes more time, which can only be compensated by increasing the bandwidth of the system.

UAVs bring a novel dynamic aspect to spectrum sharing in cellular communications. In the design of radio frequency networks, the installation of equipment and the allocation of the spectrum are traditionally carried out for individual cells. However, the movement of UAVs necessitates a more dynamic approach to cell design, taking into account factors such as UAV mobility, altitude, the number of UAVs deployed, as well as their coverage areas. AI-based spectrum sharing in UAV-enabled wireless networks is facilitated through the implementation of RL algorithms. RL models based on neural networks have been effective in various domains, but they tend to converge slowly when applied to spectrum sharing in wireless communications [91,92], a challenge also observed with FL-based models [93]. Furthermore, the reward function is critical in system optimization, making the selection of appropriate parameters and their interrelationships vital for agents to make correct decisions. The research work in the area of the UAV spectrum sharing for resource management is summarized in Table 7 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 7. AI/ML solution for UAV spectrum management.

Reference	Objective	Algorithm	Metric (s)
[91]	improve the average collision rate, throughput and the reward function	LSTM+DQN	Spectrum sensing accuracy, Channel utilization factor, mean collision rate
[92]	task offloading and channel allocation	DQN	System cost and execution time
[93]	power allocation and scheduling	FL	Convergence round of network

4.2.3. AI in Aerial Trajectory Management

Energy-efficient trajectory planning for UAVs has attracted significant research interest lately, with numerous solutions suggested for UAV-enabled wireless networks. Generally, the current strategies for energy-efficient UAV trajectory planning fall into two categories: non-ML-based methods and ML-based methods. This section delves into the ML-based methods for optimizing UAV trajectories.

In [94], the authors proposed the FL-based method for the joint optimization of the UAV position and local accuracy of the FL model and user computation and communication resources. These three problems are developed as three separate sub-problems. The proposed algorithm is compared with the fixed-altitude UAV-assisted FL ratio, performs with better learning, and reduces the system's overall energy consumption. The horizontal trajectory of UAV makes the problem non-convex, and the Successive Convex Approximation (SCA) technique is implemented to make it convex. The Dinkelbach method is applied to optimize the FL local accuracy. Finally, the Karush–Kuhn–Tucker conditions (KKTH) method is used to optimize the system bandwidth. The proposed method's performance (system cost reduction) improves as the altitude of UAV increases. In addition to this, as the bandwidth of the system increases, it supports more users and reduces the UAV energy consumption. However, this research is based on a single UAV system, and more complex multi-UAV-based scenarios need to be considered to include 3D vertical trajectory with collision avoidance and UAV transmission power to evaluate the performance of the proposed scheme.

In [95], the authors integrate DNN in UAV at MEC for communication resource allocation, model optimization, and UAV trajectory control to ensure the service latency minimization while ensuring the requirements of learning accuracy and energy consumption are met. The resulting problem is characterized as a non-convex mixed integer nonlinear programming (MINLP) problem. So the original problem is divided into three subproblems. These sub-problems are solved iteratively. By optimizing the trajectory, the UAV positions itself closer to its serving devices, thereby providing better channel conditions and reducing transmission latency. The proposed algorithm operates in polynomial time and has high complexity, making its implementation challenging, particularly when the network scale is extremely large. Moreover, the task-offloading problem is based on binary model selection variables, and each task is supported by DNN at the edge or locally. This significantly increases energy consumption limitations and computational complexities, which results in the performance deterioration of the system.

In [96], the authors maximize the sum rate of the UAV-enabled multi-cast network by jointly designing the UAV movement, RIS reflection matrix, and beam-forming design from the UAV to the users based on a multi-pass deep Q Network (BT-MP-DQN). In the proposed model, the UAV is the agent, and the beam-forming control and trajectory design are considered system actions. The movement of the UAV is discrete action, whereas the beam-forming design is continuous action. However, the UAV movement is not transformed into continuous action, which keeps this problem non-convex MINLP and the authors kept the problem non-convex. The proposed scheme is not compared with any baseline models to validate the results.

The potential for mobility that UAVs offer holds promising prospects but also introduces new challenges and technical obstacles. In UAV-assisted wireless networks, the optimization of UAV trajectories is critical, taking into account key performance metrics such as bandwidth [94], sum-rate maximization [96], energy consumption, and service latency [95]. Additionally, trajectory optimization must consider the dynamic nature and diversity of UAV types. Despite numerous studies on UAV trajectory optimization, several issues remain unresolved, including the optimization of UAV trajectories based on the mobility patterns of ground users to enhance coverage performance and the development of obstacle and collision-aware trajectory optimization for UAVs. Furthermore, the horizontal trajectory presents a non-convex problem, and it is presumed that AI-based RL techniques can manage the non-convexity. However, this assumption leads to slow convergence in the RL models. The research work in the area of UAV trajectory management is summarized in Table 8 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 8. AI/ML solution for UAV trajectory management.

Reference	Objective	Algorithm	Metric (s)
[94]	Joint UAV positioning, FL accuracy and communication resources optimization	FL	System bandwidth
[95]	Communication and computation resource allocation optimization	DNN	Latency minimization
[96]	Beamforming control and trajectory design	RL based BT-MP-DQN	Sum rate

4.2.4. AI in UAV Task Offloading and Resource Allocation

Numerous MEC-based solutions have been developed to meet the QoS requirements of data-heavy mobile applications. However, the deployment of static edge servers in isolated, mountainous, or disaster-prone regions may not be practical. In such cases, UAVs become valuable. To ensure LoS communications, UAVs can be utilized for task offloading and to improve download performance. Research typically addresses task offloading and resource allocation—such as maximizing throughput and minimizing energy consumption—concurrently. The configuration of UAV-enabled MEC systems is greatly influenced by the specific application scenario of UAV deployment. UAVs can function as relays or offloading units, temporarily handling data during high-traffic periods, which enhances system capacity as UAVs operate as base stations to meet the surge in user demand. Moreover, using UAVs as relays not only increases system capacity but also broadens coverage. In managing system load, similar to VEC, both local computing for minor data tasks and data offloading to UAVs for larger datasets are utilized. The main challenges in data offloading and resource allocation involve controlling delays and managing the relay power. The primary goals of resource allocation and data offloading are to establish connections, secure high data transfer rates, and allocate targets efficiently.

In [97], the authors address the integration of UAVs and terrestrial UE in cellular networks. The key challenge is managing inter-cell interference due to the reuse of time-frequency resource blocks. A novel approach using the first p-tier-based RB coordination criterion has been proposed. The study aims to enhance wireless transmission quality for UAVs while minimizing interference with terrestrial UEs. The goal is to minimize the UAV's ergodic outage duration (EOD). The complexity of the problem is tackled using a hybrid of Deep Double Duelling Q Network (D3QN) and Twin Delayed Deep Deterministic Policy Gradient (TD3). The proposed UAV-based system with the RL model is effective in minimizing service latency and enhancing communication quality. The study highlights the importance of practical channel modelling and advanced optimization techniques to manage the complex interference environment in cellular-connected UAV networks. However, the MINLP problem was developed with no mathematically closed-form solution,

and the authors relied on the capabilities of the RL algorithm for the optimum solution, which clearly increases the computational and time complexities.

In [98], a multi-agent DRL approach was proposed to develop an efficient resource management method for UAV-assisted IoT communication systems. The resource-management algorithm optimizes bandwidth allocation, throughput optimization, interference mitigation, and power usage management. The DRL is used with the K-means algorithm and round-robin scheduling algorithms for clustering and service request queues, respectively. The accuracy, RMSE, and testing time(s) are used as metrics to compare the proposed method with previous works, but the throughput prediction and power consumption rates are not compared with the other models and previous work. So it is difficult to assess the overall performance of the proposed algorithm.

In [99], the authors investigate the dynamic resource allocation of multiple UAV-enabled communication networks with each UAV autonomously communicating with a ground user by selecting its communicating user, power level, and sub-channel, without exchanging information with other UAVs. The long-term resource allocation problem is formulated as a stochastic game aimed at maximizing expected rewards. In this context, each UAV acts as a learning agent in the MARL model, with each resource allocation solution corresponding to an action taken by the UAVs. The reward function is based on the individual user, sub-channel, and power level decisions of UAV. However, it is considered if a UAV cannot find a user with a satisfactory QoS, it will be considered nonfunctional for the network. This makes the problem and the designed reward function very simple, such that it cannot tackle the complexities of the system. This means a complex reward function needs to be designed for the efficient UAV use.

In [100], the authors proposed a MARL approach to manage bandwidth, throughput, interference, and power usage effectively while offloading the tasks to UAV. Moreover, an actor-critic-based RL technique (A2C) solution in UAVs is implemented to offload the computational tasks of the ground users and achieve the minimum mission time. The proposed method is compared with the greedy-based method and achieves a better average response time. However, the proposed method is not compared with other advanced RL-based methods such as DDPG to compare the computational and time complexities of the algorithm. In [101], the authors probe the offloading of the task in UAV via MEC servers to minimize latency and the energy of the UAVs. Each UAV is associated with its corresponding task by keeping track of the available energy along with the optimal MEC server selection. Two Q-learning models are proposed and compared with the greedy algorithm. This study does not provide any information about the agents, states, or actions assigned in the algorithm. Moreover, no reward function is defined, and the complexity of the proposed model is not discussed either.

In [102], the authors perform task offloading to manage the resources by ensuring the energy and latency minimization for high-altitude balloon (HAB) networks. The HABs dynamically determine the optimal user association, service sequence, and task allocation to minimize the weighted sum of energy and time consumption for all users. A Support Vector Machine (SVM)-based FL algorithm is proposed to determine user association. The non-convexity is dealt with by splitting the main problem into two sub-optimization tasks: (a) optimizing the service sequence and (b) optimizing task allocation. The SVM-based global learning algorithm achieves a better accuracy rate as users vary and utility function as compared to the proposed SVM-FL algorithm. The energy consumption performance is better than baseline models as the HABs make users compute tasks locally. Moreover, the computational task time is better than other algorithms yet it is quite high and not efficient.

Given the diverse application scenarios, selecting the most suitable offloading technique is crucial for improving network throughput, bandwidth, interference [97–100], energy consumption, and latency [101,102]. In scenarios with a large number of users, network nodes such as densely populated urban areas with heterogeneous networks, deep learning approaches, or optimization-based algorithms impose higher overhead on UAVs

due to their iterative nature and longer computation and training times. Cooperative UAV-enabled hybrid algorithms are presented as a viable option, as they leverage a multi-agent system that allows for combinations of relay nodes and MEC servers. This approach enables a better selection of offloading algorithms to prevent excessive delays. The energy efficiency, flight time, and type of UAV selected for resource management and task offloading directly affect the UAVs' ability to provide long-term and viable alternatives to the MEC server. Moreover, the binary offloading and allocation problem is mostly formulated where either the task is computed at the device level or completely offloaded to the UAVs. This approach limits the implementation of UAVs as most of the time, due to energy constraints of UAVs, they cannot compute the complete task. The research work in the area of UAV-based task offloading and resource allocation for resource management is summarized in Table 9 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 9. AI/ML solution for UAV task offloading and resource allocation.

Reference	Objective	Algorithm	Metric (s)
[97]	inter-cell interference to minimize latency and communication quality	D3QN	ergodic outage duration
[98]	optimizes bandwidth allocation, throughput optimization, interference mitigation, and power usage management	MADRL	accuracy, RMSE and testing time(s)
[99]	dynamic resource allocation of multiple UAV-enabled communication networks	MARL	average reward
[100]	manage bandwidth, throughput, interference, and power usage effectively	MARL	average response time
[101]	minimize latency and UAV energy	Q-Learning	processing time and energy
[102]	energy and latency minimization	SVM based FL	accuracy, task completion time and energy consumption

4.3. AI for Resource Management in UAV-IoV Networks

The heterogeneity of vehicular networks and their highly dynamic nature with fast-moving wireless nodes have made them more complex and demand new requirements for networking algorithms that can meet the stringent network control and resource-allocation demands such as efficient spectrum sharing, transmission power maximization, and computational resource management to minimize the energy requirements of UAVs and vehicles' local computation. UAV-IoV networks are three-dimensional and contrast with terrestrial networks, and the UAV-BS itself moves with the vehicles on the roads. Therefore, traditional optimization techniques are unable to capture complex patterns. The resource management in UAV-IoV is divided into radio resource allocation and computational resource management. The radio resource allocation is further divided into spectrum and channel access optimization. The main goal of radio resource management is to limit channel interference, power usage, and network congestion. The computation resource management includes service, task, and traffic offloading in MEC, where the edge cloud nodes are located in BSs and/or UAVs. This decentralization of the system generates faster response times compared to the central deployments. In this section, we review AI-based resource allocation research conducted in UAV-IoV networks.

4.3.1. AI Deployment of UAV-IoV Systems

The integration of UAVs and vehicles within an AI-based IoV network enabled by UAVs is an under-researched area. Most studies focus on UAVs with static IoT users, cellular

BS, vehicles, and RSUs. However, considering the high mobility of both UAVs and vehicles, vehicle clustering on roads and UAV deployment in the aerial network become critical due to the rapidly changing channel conditions between UAVs and vehicles. Additionally, with vehicles traveling at varying speeds, maintaining favorable channel conditions to ensure good QoS is essential, necessitating a mechanism for adequate connection time.

In [103], the authors proposed an FL-based approach to the development of IoV-based applications. The authors used the Gale–Shapley algorithm to match the lowest-costing UAV to each sub-region. The UAV performs the local training. Based on the transversal and transmission cost function, the multi-dimensional node-coverage cost is converted into a single-dimensional node coverage. The simulation results show that the lowest marginal cost of node coverage for a UAV is assigned to each sub-region for task completion. The UAV energy constraint has not been considered as much as the effect of the flight time on the node coverage. Moreover, the proposed technique is not compared with previous works or any other baseline model to provide comprehensive analysis in terms of throughput maximization, energy consumption, and computational complexity.

In [104], the authors deployed UAVs as relays to improve the communication efficiency between the model owner/server and the workers/vehicles. The paper combines auction-integration (AI) formations to integrate UAVs into groups of IoV elements with the target of achieving the total revenue maximization of a single UAV. The algorithm becomes more complex as the number of UAVs is increased, which in turn exponentially increases the number of partition sets of UAVs that need to be found. So the model is not affected by the change in the number of vehicles, but at the same time, if the overall size of the cell increases, it affects the communication efficiency of the proposed model, and the authors did not tackle this issue in this study.

A vast amount of research has been conducted on the deployment of vehicles and UAVs individually using AI. Yet the deployment of UAVs in relation to the distribution of vehicles on roads remains unexplored. The energy constraints, the relative speed of UAVs to vehicles on the road, and the UAVs' brief flight duration present significant challenges in managing communication resources within UAV-based IoV networks. The research work in the area of UAV and vehicular deployment for resource management in UAV-assisted IoV networks is summarized in Table 10 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 10. AI/ML solution for UAV-IoV deployment.

Reference	Objective	Algorithm	Metric (s)
[103]	target sensing region to fulfill a time-sensitive task	FL	UAV utility
[104]	improve the communication efficiency	FL	communication time

4.3.2. AI in Resource Allocation and Task Offloading in UAV-IoV Networks

In VEC networks, vehicles can offload computationally heavy applications to vehicular edge servers, like RSUs, for processing. This offloading leads to decreased processing time and reduced energy use. However, in densely populated areas or on busy roads, RSUs can become overwhelmed, and their performance deteriorates as vehicle density grows. To address this, UAVs have been integrated into IoV networks to assist with the computational burden of overloaded RSUs, offering improved resource distribution and task offloading capabilities through UAV-based edge computing.

In [105], the vehicular task-offloading optimization problem is dealt with by jointly considering the task offloading, resource allocation, and security assurance. This is an MINLP, non-convex, and non-deterministic polynomials (NP) problem. Therefore, this problem is divided into two separate problems, and finally, the iterative algorithm called LBTO is proposed. LBTO decides if a certain MEC is selected depending on the load of the MECs and uses the Lagrangian dual decomposition for the optimized offloading ratio

and the computation resource. The task to be processed is selected based on the size of the task, the computing resources required to execute a task, the task's allowed latency, and the ratio of the offloaded task at UAV/MEC or locally to the total task. The functionality of the proposed algorithm is not explained in detail. The proposed algorithm provides a better task offloading ratio and delay than the other algorithms. However, this research considers the UAVs to be fixed, and that is why it completely ignores their energy consumption during mobility in the objective function. This clearly ignores the flight energy used by UAVs, which is a major source of energy consumption and affects the ability of UAVs to support the computation task processing.

In [106], the authors proposed a mechanism for energy harvesting by UAVs from BS and vehicles using wireless power transfer (WPT) and simultaneous wireless information and power transfer (SWIPT) techniques, respectively. Maximum data offloading to the UAV is the main goal of this research, which in turn maximizes the throughput of the system by jointly maximizing the computational resource offloading, the amount of the task to be offloaded, and the speed of the UAVs. The DRL-based resource allocation and speed optimization (DRL-RASO) model is adopted. The state space includes locations of UAVs and vehicles, the current on-board energy of UAVs, and the speed of UAVs. The action space includes the resources allocated to vehicles, the tasks to be offloaded to the UAV, and the speed of the UAV. The reward function of the proposed model does not converge as fast as Dueling-DQN, and a lot of fluctuations can be observed throughout the process. Clearly, the actions generated by the classical Dueling-DQN are discrete, resulting in a significantly smaller overall action space compared to the continuous action space of the proposed algorithm.

In [107], a UAV-based vehicular network is built to deal with caching and computing problems in addition to BS. The energy minimization is achieved by combining the cache refreshing optimization, computation unloading, and status age updates. The on-line decision-making is performed using DDPG. The BS decides if the cache needs to be refreshed, if the task has to be executed, and what the bandwidth distribution should be. The total energy consumption is the reward function. The learning performance of the proposed model is compared with the traditional DDPG algorithm in terms of the convergence rate. Then, the energy consumption for four benchmarks, namely random refreshing, random offloading, popular refreshing, and equal bandwidth, is calculated. The proposed model outperforms DDPG in terms of system energy consumption and computational capabilities of the UAV MEC server. But the authors do not report any results obtained using the DDPG model.

The authors in [108] proposed a secure bandwidth-allocation scheme based on game theory for IoV-assisted UAV communication systems. To allocate the limited safe bandwidth, based on the real-time feedback of each UAV, an optimal decision search algorithm based on gradient descent to achieve Stackelberg equilibrium is proposed. The proposed scheme achieves a better throughput of about 95% compared to other models, but the authors do not provide any data to strengthen their claim about privacy and secured bandwidth allocation.

In [109], the authors proposed a model-free Q network to select the best UAV advice with the lowest stalling time. The problem is dealt with as a binary download problem, wherein, if the UAV is positioned in the requested vehicle section, it fulfills the request and the UAV otherwise drops it. The results show that the proposed system takes longer to converge and the reward function fluctuates throughout the training. The complexity of the algorithm is not discussed, which is a crucial factor in evaluating the performance. As the number of vehicles sending download requests to the UAVs grows (the state space grows), the action (support download or drop the request) space grows as well. Therefore, the neural network used with the RL algorithm takes more time to converge. This paper is limited in its scope as the limitation of UAV server capacity, speed, and vehicle speed are not considered.

AI is pivotal in integrating complex mobile UAV and IoV networks to facilitate task offloading and resource allocation. The primary research focus is on addressing the non-convex and mixed-integer challenges presented by UAV and IoV networks [105]. This includes offloading tasks to UAV MECs to alleviate the load on RSUs while maximizing throughput [106,108], reducing latency [105], and minimizing energy consumption [107]. Reinforcement Learning algorithms are employed to optimize the system. However, a significant challenge lies in designing an accurate reward function that will guide the agents' future actions. Additionally, the high mobility factor of vehicles is often overlooked in the design of UAV-enabled IoV networks, simplifying channel conditions and offloading decisions. The research work in the area of UAV-assisted IoV networks for task offloading and resource allocation is summarized in Table 11 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 11. AI/ML solution for UAV-IoV task offloading and resource allocation.

Reference	Objective	Algorithm	Metric (s)
[105]	task offloading, resource allocation and the security assurance	LBTO	task offloading ratio and delay
[106]	maximum data offloading to the UAV	DRL-RASO	throughput maximization
[107]	dynamic resource allocation of multiple UAV-enabled communication networks	DDPG	energy minimization and convergence rate
[108]	secure bandwidth allocation	game theory	throughput
[109]	best UAV advice with the lowest stalling time	RL	reward function convergence

4.3.3. AI in Trajectory Management of UAV-IoV Networks

UAVs provide the adaptability to modify their positions based on real-time traffic needs, guaranteeing network connectivity in situations where ground networks are compromised or non-existent due to geographical constraints. In dynamic vehicular networks with varying vehicle arrival patterns, the deployment of multiple autonomously controlled UAVs is essential for collaboratively sustaining network coverage and adapting to fluctuating traffic dynamics. Extensive research has been conducted on the optimal positioning and path planning of UAVs to overcome these challenges.

In [110] the authors proposed a Markov Decision Process (MDP)-based model for UAVs to optimize the UAVs' trajectories to minimize the number of UAVs that serve vehicles within the highway segment under the mobility of UAVs and vehicle constraints, as well as the UAVs' energy budget constraint. The actor-critic algorithm learns the environment. The problem under consideration is an MINLP and non-convex problem. The DRL model is used to learn this underlying non-linearity and non-convexity optimally. The model inputs are the residual energy of each UAV, the number and position of vehicles, the positions of UAVs with respect to ground level, etc. The UAVs' traveling distance is taken as the action. The penalty on the network incurred if the UAV does not provide coverage to a vehicle, a new UAV is deployed, there remaining energy for each UAV, or the UAV goes outside the designated path. The reward function converges quickly and remains smooth on average. As the requirement for the minimum data rate varies, more UAVs are required to fulfill the demand. The UAVs change trajectory to reduce the distance from the vehicles to meet the requirements. This study considers all the important aspects of UAV-based IoV communication networks, and it can be concluded that the actor-critic DRL model can produce stable and satisfactory results if the appropriate reward function, action, and state spaces are chosen carefully.

In [111], the authors focus on bandwidth allocation, location control deployment, and the trajectory of UAVs for average communication channel capacity (throughput) maxi-

mization to enable the UAV to process more data with edge computing. They propose an actor–critic mixing network (AC-Mix) and a multi-attentive DDPG (MA2DDPG) network. The AC-Mix is the combination of Qmix (it relies on the Q function and does not deal with continuous values) with the actor–critic framework. The reward function is based on the addition of four individual reward functions, namely achievable capacity, low-SNR penalty, collision penalty, and out-of-bounds penalty. The proposed model converges faster than the comparative models as the critic uses the local information as input.

In UAV-enabled IoV networks, addressing the continuous UAV trajectory optimization problem is analytically challenging due to the need to determine an infinite number of optimization variables, specifically the UAV locations. Furthermore, in vehicular networks, no current framework can ascertain the minimum number of UAVs required to serve vehicles on a specific highway segment in a high-mobility scenario while complying with the UAVs' energy constraints and ensuring a satisfactory QoS for each vehicle. Traditional coverage approaches often presuppose stationary users and depend on complete environmental knowledge, including real-time user locations, to produce accurate results. However, this assumption does not hold in dynamic environments like vehicular networks, where users, such as vehicles, may travel at varying speeds, thus invalidating the premise of global network knowledge. The research work on UAV and vehicular trajectory control and optimization for resource management in the area of UAV-assisted IoV networks is summarized in Table 12 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 12. AI/ML solution for UAV-IoV trajectory management.

Reference	Objective	Algorithm	Metric (s)
[110]	optimize the UAVs' trajectories to minimize the number of UAVs	DRL	average coverage and maximum performance
[111]	average communication channel capacity (throughput) maximization	MA2DDPG	reward functions based on capacity, low-SNR penalty, collision penalty, and out-of-bounds penalty

4.4. Joint Resource Management Metrics in UAV-Assisted IoV Networks

Figure 6 summarizes the joint resource metrics considered in the literature for effective system performance. In [72], the authors jointly optimized the vehicle position and cache allocation vehicular network using a supervised-learning-based joint time series method. The vehicle movement is predicted using LSTM. The caching strategy is obtained using the heuristic ϵ_n greedy process. In [62], the authors optimize the spectrum and power allocation using the Q-learning-based reinforcement learning method and achieve an improved sum capacity of V2I links and payload delivery rate of V2V links. The cumulative reward guarantees the delivery of a large amount of V2V data until the payload is ended. The resource-sharing algorithm is trained offline as it is a computationally intensive algorithm. In [65], the authors proposed a reinforcement-learning-based scheme in which the BS acts as an agent and VANET is the environment, and the end goal is to jointly optimize the cluster head transmission power and maximize network energy efficiency under given latency limitations. First, vehicles are divided into clusters and the cluster head communicates with the BS. The BS provides the data requested by the cluster head, and if BS cannot handle the amount of data requested, it offloads the data-processing tasks to the cloud. In [58,67], the authors proposed a reinforcement-learning-based method with a V2V link as an agent to select the frequency band and transmission power level that cause minimal interference to both V2I links and other V2V links, ensuring sufficient resources are preserved to satisfy latency constraints. In this regard, the reward function includes the capacity of the V2I links, the capacity of the V2V links, and the latency condition. The latency condition is introduced as a penalty. Deep Q-learning is used in the resource allocation scenario, and after identifying the optimal policy through training, it is utilized to

choose spectrum bands and transmission power levels for V2V links, aiming to maximize the overall capacity while maintaining latency constraints for V2V links.

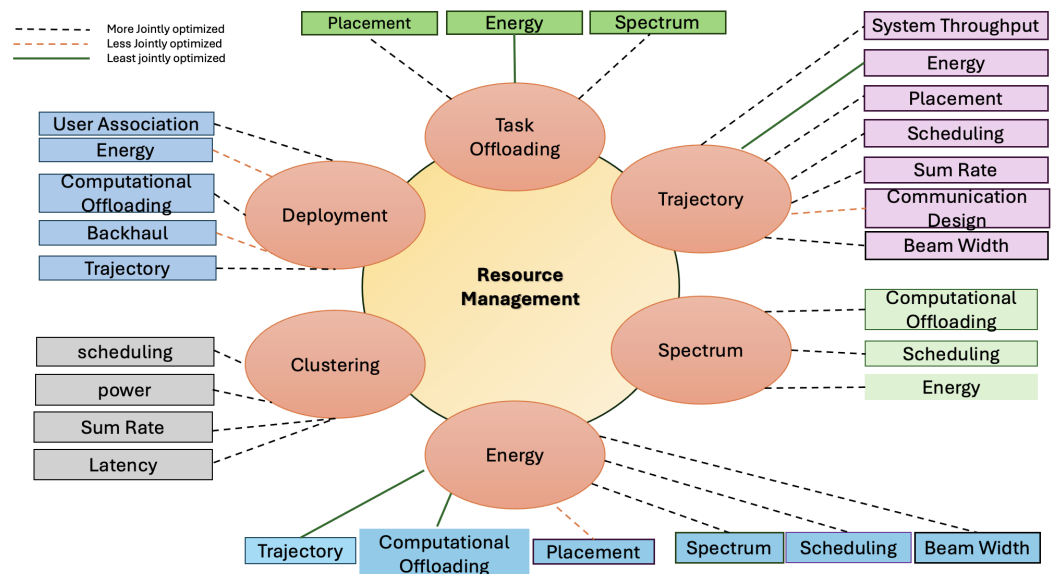


Figure 6. Joint resource management metrics in UAV-assisted IoV Networks.

In [89], the authors achieve minimum UAV energy consumption while clustering the network nodes using the K-mean algorithm to maximize the data collection from the nodes. A multi-objective joint DDPG algorithm is proposed for the multi-objective control policy of UAVs by jointly optimizing the UAV flight decision, hovering time slot, and UAV launch power. In [94], the effect of changing UAV altitude on the communication area, a joint optimization of UAV placement and computation, and communication resources is proposed. The federated learning algorithm is used, which makes this optimization problem non-convex. To handle this issue, this one problem is decomposed into three different optimization sub-problems namely UAV horizontal placement, local accuracy, and computation and communication resources. In [95], the authors achieve the trajectory control of UAV and computational and communication resource optimization using DNN under the energy consumption, latency, computation, and communication resource constraints. In [96], the problem of maximizing the sum rate involves formulating a solution that coordinates the movement of the UAV, RIS reflection matrix, and the design of beamforming from the UAV to users. This paper introduces a novel approach called the Beamforming control and Trajectory design algorithm, which utilizes a Multi-Pass Deep QNetwork (BT-MP-DQN) for efficient optimization. In [91], the authors proposed a LSTM+DQN-based algorithm to jointly optimize the channel allocation and time slot allocation in UAVs based on the priority of the task. In [92], the authors introduced UAV as a relay and edge computing node to process tasks offloaded by the vehicles. An optimal available channel allocation based on the OFDMA scheme is proposed. In [107], the authors proposed an RL (DDPG)-based energy minimization mechanism by jointly considering cache refreshing, computation offloading, and aging of the status updates. In [111], the authors used UAVs as edge computing devices to accommodate vehicles. To achieve this goal, the authors focused on bandwidth allocation and UAV trajectory control to maximize the system's communication capacity.

Research in UAV-based IoV networks is still emerging, with many unresolved issues. The primary research focus has been on minimizing system energy and optimizing energy harvesting, caching, and bandwidth allocation. However, this focus often overlooks latency requirements, UAV altitude adjustments, and variations in vehicular node density. Reinforcement learning is commonly used for resource management, yet state-of-the-art RL models primarily address system energy and bandwidth allocation. The scarcity of datasets

for UAV-based IoV networks hinders the use of ML and DL models for resource management and limits the exploration of these AI models' capabilities. Moreover, DL techniques remain under-explored due to UAVs' limited power and processing resources. Additionally, critical issues of security and data privacy are frequently neglected. UAV communications often utilize unencrypted and unauthenticated channels, exposing them to cyber threats. Federated learning could significantly enhance security and privacy by enabling ML models to be trained on data locally without being transferred to a cloud server.

5. AI-Based Routing in UAV and IoV Networks

In this section, we cover the research contribution in the area of routing protocols proposed and designed for IoV, UAV, and UAV-assisted IoV networks using AI.

5.1. Classification of Routing Protocols

There are three types of routing protocols designed for the UAV-IoV networks based on position, topology, and AI-enabled routing. These routing protocols are further divided into different categories, as shown in Figure 7. A detailed discussion about the routing protocols is provided in this section below.

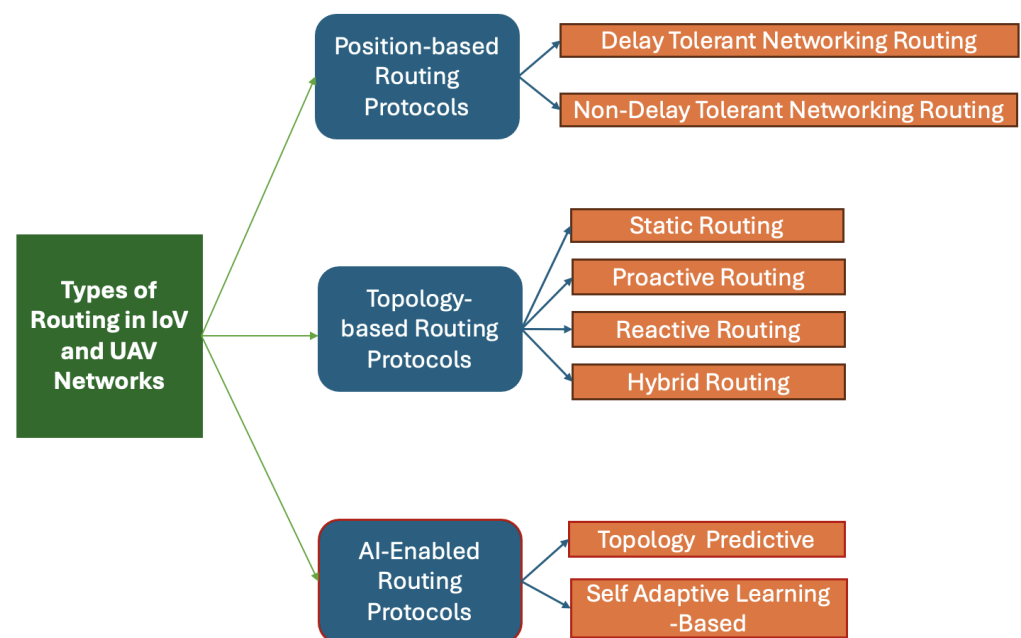


Figure 7. Classification of routing protocols in UAV-IoV Networks.

5.1.1. Position-Based Routing Schemes

These methodologies leverage the geographical data of nodes. Therefore, each node interfaces with a positioning system, such as the Global Positioning System (GPS), to access its spatial information whenever needed [112,113]. These routing techniques do not necessitate complete network information and rely on local data, enhancing communication efficiency, reducing bandwidth usage, and conserving energy. Consequently, they are particularly suitable for highly dynamic networks like VANET. These approaches are typically categorized into two groups:

- **Delay-Tolerant Network Routing:** These methods effectively address the challenges arising from frequent disconnections in VANETs, which often result in broken paths to the destination node. Typically, these approaches employ the store–carry–forward technique when a node is unable to establish a routing path to other nodes [114,115]. While this technique significantly reduces communication overhead by eliminating

the need for additional control packets, it does introduce delays in the data transfer process [116,117].

- **Non-Delay-Tolerant Networking (non-DTN) Routing methods:** These protocols are designed for use in networks with high connectivity, where node density is relatively dense. However, if network connectivity cannot be guaranteed, the performance of these protocols may be compromised. They typically employ a greedy forwarding technique for data transmission [118], where transmitters send data packets to the neighbor closest to the destination. If the sender cannot find a neighbor closer to the destination than itself, then the data delivery process may fail, necessitating the use of a recovery strategy to manage this situation. These methods demonstrate good performance in high-density networks, exhibiting low communication overhead, high scalability, and low memory requirements. However, a significant challenge lies in obtaining accurate location information. If node locations are unavailable or inaccurately calculated, these protocols may exhibit weak performance. Moreover, since all nodes are equipped with GPS in these methods, significant bandwidth is required.

5.1.2. Topology-Based Routing Schemes

In these approaches, topological information about nodes is utilized for transmitting data packets within the network [118–120]. They establish a suitable path before initiating the data-transfer procedure. Topology-based routing methods are typically classified into four groups:

- **Static routing protocol:** Static routing protocols feature fixed and non-modifiable routing tables, which are primarily suited for scenarios with stable topologies and no task updates. However, traditional static routing protocols have limited applications in UAV swarm systems due to their lack of fault tolerance and adaptability to dynamic environments. Three static routing protocols are load carry and deliver routing (LCDR) [121], used for centralized communication architecture, data-centering routing (DCR) for one-to-one data transmission requirements in IoV and UAV environments, and multilevel hierarchical routing (MLHR) to solve the scalability problem in UAV and vehicular networks.
- **Proactive routing methods:** Also known as table-driven protocols, these approaches involve each vehicle continuously transmitting the latest routing information to other vehicles, regardless of whether they have data packets to send. The routing information is stored in the routing tables of vehicles and is regularly refreshed and shared with network nodes. Proactive routing is not well suited for VANETs due to their limited ability to respond effectively to frequent topological changes, leading to high route breakage. Currently, the most widely used proactive routing protocols include the Optimized Link State Routing (OLSR) protocol with flat topology, the Destination Sequenced Distance Vector (DSDV) protocol to provide nodes for every path in the network from source to destination, and their respective variations [122].
- **Reactive routing methods:** These approaches operate on demand. When a vehicle has a data packet to deliver to a destination and there is no existing path for this purpose, it initiates the route-discovery process. In these protocols, vehicles maintain routing information solely about valid paths. Consequently, a path maintenance system verifies valid paths and eliminates invalid ones. Upon updating the network topology, failed paths are removed, and the route-discovery process restarts. Reactive routing protocols are more efficient in terms of bandwidth consumption compared to proactive routing methods, as routing tables are updated only as needed. The main reactive protocols are Dynamic Source Routing (DSR) and Ad hoc On-Demand Distance Vector (AODV) [122].
- **Hybrid routing protocols:** Combining proactive and reactive approaches, hybrid routing aims to mitigate their respective weaknesses. This method reduces communication overhead compared to proactive routing protocols and enhances delay in the path

discovery process compared to reactive routing schemes. Hybrid routing protocols are particularly suitable for large-scale networks. The Zone Routing Protocol (ZRP) and Temporarily Ordered Routing Algorithm (TORA) are two major protocols that represent hybrid routing protocols [122].

5.1.3. AI-Enabled Routing Protocols

AI-enabled routing protocols leverage the learning capabilities of Machine Learning (ML) algorithms to select optimal route paths based on a precise understanding of network topology, channel conditions, user behavior, traffic mobility, and other factors. These algorithms integrate networking and AI research to realize advanced networking, particularly for dynamic UAV and IoV networks.

- **Topology predictive protocol:** The primary characteristic of topology-predictive routing protocols lies in their utilization of Machine Learning (ML) algorithms to forecast node motion trajectories. These trajectories, serving as an approximation of the network topology, are integrated into the path selection mechanism, particularly when the communication range of nodes is known.
- **Self-adaptive learning-based routing protocols:** Most learning-based routing protocols employ RL to make routing decisions through the continual and online learning of the environment and their decision consequences on desired performance metrics such as delay, throughput, energy efficiency, and fairness. RL-based algorithms offer a significant advantage due to their abstract formulation, which grants independence from topology prediction and channel estimation, thanks to the concept of learning from experience. The concept of RL for optimized routing is depicted in Figure 7. Initially, the scenario is represented by state S_1 , where the node or agent A_1 has two candidate neighbors, A_2 and A_3 , to send its packet to. Subsequently, a choice is made between actions a_1 or a_2 based on the expected reward for each action a at state s , defined as $Q(s, a)$. Upon selecting the appropriate action, agent A_1 receives an immediate reward from the environment, r_1 or r_2 . This process repeats in a new state s_2 , where decisions are made based on the new environmental conditions and the learned policy in terms of action–reward relations. The ultimate objective is to identify an optimal policy wherein the cumulative reward over time is maximized by assigning optimal actions to each state [123]. RL-based routing was initially introduced in [124], where Q-Routing treated packet forwarding as an application of Q-learning. This method exhibited superior performance compared to a non-adaptive algorithm based on pre-computed shortest paths [125]. The essence of Q-Routing lies in evaluating the impact of routing strategies on desired performance metrics by exploring different paths in the exploration phase and utilizing the best paths discovered in the exploitation phase. While exploration imposes overhead on the system, it is crucial for identifying newly optimal paths, especially when the network topology undergoes significant changes. An inherent challenge is adaptively resolving the trade-off between exploration and exploitation times to accommodate the dynamicity of the network topology.

In the next section, the topology-based and self-adaptive learning-based routing protocols in UAV and IoV networks are summarized to understand the evolution of these protocols over time.

5.2. AI for Routing in IoV Networks

In opportunistic networks, node selection poses a crucial challenge due to nodes lacking information about the state of other nodes. Furthermore, in IoV, traditional routing protocols fall short of achieving optimal performance. To address these challenges, the authors introduced a Machine-Learning-based multi-copy routing algorithm called iPRoPHET (Improved PRoPHET) in their work [126]. iPRoPHET leverages dynamically changing contextual information of nodes and the delivery probability of PRoPHET for effective message transfer. Employing a random forest, iPRoPHET classifies nodes as reliable or non-reliable forwarders based on contextual information provided during each

routing decision. The training data are derived from simulations. The proposed model undergoes evaluation using metrics such as delivery probability, hop count, overhead ratio, and latency, demonstrating performance on par with similar multi-copy routing algorithms. The comparison of the proposed scheme is not provided with other proven state-of-the-art ML algorithms. Moreover, in terms of latency, overhead ratio, and hop count, the proposed algorithm does not perform better than some of the baseline models.

In [127], the authors proposed a stochastic chaos-based adaptive routing with prediction (SCARP) to predict traffic flow using DL networks to suggest a node-discovery routing principle. The connectivity loss and delay are minimized and guarantee a secure data transmission between vehicles. In this research, the region of Puducherry U.T., India, is selected for traffic data collection. The simulation software, namely Simulation of Urban MObility (SUMO) and Objective Modular Network Testbed in C++(OMNET++), are used to create traffic and network scenarios, respectively. The metrics of accuracy, PDR, delay, and sensitivity are used to compare the proposed method with existing state-of-the-art routing algorithms. The metrics used for prediction are accuracy, precision, and recall. The study compares its results with previous studies with the integration of chaotic encryption in data transmission during routing in detail and reports better results even with a higher probability of attacks.

In [128], a Q-learning-based geographical routing scheme with intersection-based V2X routing (IV2XQ) is introduced. First, the best road segment at the intersections for routing is selected using Q-learning at intersections. Then best relay node selection is performed using a greedy routing strategy. The central server is the agent that uses historical traffic data to select the optimal path. The environment, which is our entire network, rewards the agent if it takes the right action and chooses the correct road segment to forward a data packet to. It is reported that the proposed scheme increased the PDR, minimized the communication overhead and latency, and considerably controlled the network congestion. The proposed algorithm was not compared with any other state-of-the-art RL algorithms, which also include the Q-tables and learning.

In [129], an RL-based routing (best two hops) and context-aware edge node selection scheme to forward packets scheme (CEPF) is proposed. Both unicast and broadcast communications are supported by CEPF. This routing protocol reduces the forward nodes and increases the resource efficiency. Decentralized fuzzy logic is implemented to select the edge nodes based on vehicular velocity, mobile nodes traveling in the same direction, and communication link conditions. The edge node is the vehicle with highest node score. For the route discovery operation, RL is used, in which each packet is the agent and the action that the agent takes is the selection of the next-hop node. The reward is awarded when the source node is one hop to the destination node. The packet delivery ratio deteriorates as the number of flows increase resulting in a large number of hops.

In [130], the authors integrate the RL and fuzzy logic and propose a reinforcement routing protocol named RRPV. A DynaQ technique is implemented on the fuzzy logic to build the model. The link stability and connection quality are two inputs for the fuzzy-logic-based system. The fuzzy system determines the link quality, and this result is fed as the state transition probability in MDP. In the RL process, the vehicles are agents, and each agent has two states, namely F to send a packet and D to deliver packets to adjacent vehicles. A hello message delivery to the neighboring vehicle is the action of the vehicles. Moreover, the link condition and the Euclidean distance of two neighboring vehicles define the reward function. The model-based Q-learning and model-free approaches are used. As the speed of the vehicles increases, the transmission delay increases as the link quality deteriorates so the proposed model does not give a satisfactory performance. However, as the number of nodes increases, it provides more links, and the packet delivery ratio improves. The proposed MARL model with FL is not compared with any other RL-based model, and the computational and convergence analysis of the model is not provided either.

In [131], the authors introduced traffic-aware routing protocol based on Q-learning (QTAR). It contains two routing algorithms to send data packets between vehicles (V2V

Q-learning) and between RSUs (R2R Q-learning). Vehicles broadcast HelloV2V messages containing their velocity and location-related information, and the RSUs exchange HelloR2R messages with each other. The reward function is formulated by including the link quality, link expiration time, and the delay. So the reward is high for selecting the next hop link with good quality, a long survival time, and a short delay. The proposed technique performs better in terms of the packet delivery ratio but the end-to-end delay is at par with other baseline models.

The authors of [132] propose a routing protocol named RLRC for clustered networks based on the K-harmonic means (KHM) clustering to assign vehicles to different cluster and RL to exchange data between two CHs. In RLRC, a hello message is used to share the vehicle velocity and position with neighboring vehicles. In this process, each node behaves as an agent, the state set is defined as neighboring CHs, and a next-hop CH selection is the action taken by the agents. The reward function is based on the link quality parameter. If the current node is the neighbor of the destination node, the reward is 1 and otherwise 0. The bandwidth availability and connection duration are used as the evaluation indicators of link status. Moreover, the final Q-value is based on the values of hop counts, link utility, and bandwidth. The proposed model is compared with baseline routing protocols but not with other RL-based routing protocols.

In [133], the authors proposed a Q-learning-based routing scheme called a reliable self-adaptive routing scheme (RSAR). The vehicles are agents. The action is a beacon message including the vehicle speed, location, and Q value, sent to the next vehicle. Moreover, the decentralized learning process is adopted with the number of hops, bandwidth, and link reliability as learning parameters. The RSAR finds the fittest relay vehicular node and solves the network segmentation problem. The proposed model does not perform better than other Q-learning-based and classical routing protocols in terms of average route length (number of hops to reach destination). The Q-learning-based AODV protocol achieves almost the same results as the proposed model in terms of the packet delivery ratio.

In [134], the authors proposed a routing technique that allows the central server and vehicles and RSUs network nodes to access the updated traffic information based on intersection-based Q-learning (IRQ). The global traffic view is obtained by IRQ, as well as by the central server to form a routing solution. Here, the central server behaves as an agent. The central server is also responsible for network congestion control in the routes. The IRQ uses a greedy routing approach in V2V and V2R routing decisions, where, for V2V routing, the vehicle closest to the target is chosen to forward the data packet and in the V2I scenario, the RSU located at the intersection delivers data packets to the corresponding road section. If there is no vehicle available, RSU waits and holds the packet until it finds a vehicle to relay information. The performance of the proposed model is compared with the IV2XQ [128], Q-learning, and grid-based routing protocol (QGrid), as well as Greedy Perimeter Stateless Routing (GPSR). The reward function is based on the vehicle density, the average connection time, and the average delay in the current road segment. The IV2XQ attains a better overhead ratio than the proposed IRQ as the proposed model does not use historic traffic information. In addition, the average hop count of the QGrid is better than that of the IRQ.

In [135], a routing protocol based on Q-learning and a fuzzy-based hierarchy (QFHR) is proposed. The routing algorithm is capable of carrying out traffic pattern recognition and routing between intersections and at road sections. The RSUs are equipped with Q-learning to find multiple routing paths. Moreover, the vehicles use the greedy technique to find the best-fitting path in each road section. A fuzzy solution works as the alternate for route recovery if the main algorithm fails and selects the next node. The proposed scheme is compared with IRQ [134], IV2XQ [128], QGrid, and GPSR. The proposed model outperforms other protocols in terms of packet delivery ratio and the average hop count. However, for the overhead ratio, IV2XQ and IRQ perform better than the proposed QFHR protocol. One reason is that the vehicular clustering is not considered, so it contributes to the overhead.

In summary, the incorporation of AI into routing protocols predominantly utilizes RL models. While RL has proven quite effective in making routing decisions, it is not without its limitations. The extensive state and action sets can slow the convergence rate and add delays to the routing process. Future research should concentrate on refining state and action spaces according to specific criteria. Currently, most studies employ a Q-table, and as the state and action sets expand, so does the Q-table's dimensionality, necessitating more memory and consequently increasing system latency. Thus, future studies should address Q-table management. Additionally, in RL-based routing, the dynamic adjustment of learning parameters is essential to balance exploration and exploitation, a factor that warrants further attention. Furthermore, predictive RL approaches should be explored, as accurately forecasting Q-values is vital for the RL algorithm to make more precise routing decisions. The research work in the area of AI-based IoV network routing for resource management is summarized in Table 13 based on the objectives of the research, the algorithm design, and the metrics used to evaluate the performance of the proposed algorithm.

Table 13. AI/ML- based routing solution for IoV networks.

Reference	Objective	Algorithm	Metric (s)
[126]	classification of nodes as reliable or non-reliable forwarders based on contextual information	Improved PROPHET	delivery probability, hop count, overhead ratio, and latency
[127]	predict traffic flow	SCARP, SUMO and OMNET++	accuracy, PDR, delay and sensitivity
[128]	best road segment and relay node selection	IV2XQ	PDR, communication overhead and latency
[129]	context-aware edge node selection	RL based CEPF	PDR
[130]	Efficient packet delivery and reception from the adjacent vehicles	RL and fuzzy logic (RRPV)	PDR and link quality
[131]	send data packets between vehicles and RSUs	Q-learning based QTAR	PDR and end-to-end delay
[132]	Divide vehicles in clusters and communication between CHs	RLRC	hop counts, link utility and bandwidth
[133]	Send beacon message including vehicle speed, location, and Q value to the next vehicle	Q-learning based RSAR	average route length
[134]	Access to the updated traffic information for the central server, vehicles, and RSUs	Q-learning (IRQ)	overhead ratio and average hop count
[135]	traffic pattern recognition, routing between intersections, and at road sections	Q-learning based(QFHR)	PDR and average hope count

5.3. AI for Routing in UAV Networks

In [136], the management of multiple cooperative UAVs is addressed. The routing problem in this system is divided into two stages: initial planning and THE routing solution. In the initial planning stage, regions to be visited are grouped into clusters based on the distance criterion (FCM algorithm), with each cluster assigned to a UAV. The route-solving stage determines the best route for each agent, considering the clusters from the initial planning stage and a variant of the Orienteering Problem. The Transformer deep learning architecture is employed to solve the Orienteering Problem with shared regions, coupled with a DRL framework. The proposed model is evaluated using multiple OP-MP-TN datasets under various environmental conditions, demonstrating its superiority over state-of-the-art models in cooperative and non-cooperative scenarios.

In the work presented by [137], UAV location optimization and relay path planning are jointly achieved using a graph neural network based on the RL (RGNN) algorithm.

The proposed model exhibits significantly lower time complexity compared to traditional optimization methods. The location GNN (LGNN) optimizes UAV locations, and the RGNN selects the optimal relay path based on information provided by the LGNN. The method outperforms the Bellman–Ford approach in terms of the data rate achieved and time complexity. The proposed model achieves the same data rate compared to Bellman–Ford but the time complexity of the proposed model is very low as compared to Bellman–Ford because of the parallel computing in RGNN.

The study conducted by [138] introduces a novel routing protocol based on ant behavior routing, enhancing end-to-end security through data encryption using the Pheromone update process. Experiments conducted in Network Simulator-2 show that AntHocNet performs well in terms of packet drop rate, throughput, and bandwidth utilization, achieving significant optimizations compared to other routing techniques.

In [139], collision-free routing policies for UAVs are designed using MARL. The authors propose a multi-resolution, multi-agent, mean-field RL algorithm named 3M-RL for UAV flight planning. Each UAV makes decisions based on local observations without direct communication with other UAVs. A UAV does not know the decision and condition of the other UAVs while taking action. The routing policy is trained using a CNN-based actor-critic neural network with multi-resolution observations, demonstrating effectiveness in various complex scenarios in both 2D and 3D space, but as the grid size increases, the CNN algorithm performance deteriorates. The environment is discrete in time, continuous in state space, and discrete in action space. This makes the problem MINLP. This issue is not dealt with in the study, and the performance of the propose technique is not compared with any other RL based on the classical routing technique.

The predictive ad hoc routing combined with RL and the trajectory knowledge protocol (PARRoT) is introduced by [140]. This protocol aims to achieve lower latency and high robustness by predicting future node positions and sharing information with adjacent nodes. The PARRoT separates networking from path planning, enhancing the overall system efficiency.

In [141], fuzzy logic is employed to identify adjacent nodes in real time, while RL is used to reduce the number of hops in a routing algorithm named the Fuzzy Logic Reinforcement Learning-based Routing Algorithm (FLRLR). The FLRLR reduces the average number of simulation hops and ensures higher link connectivity, showing comparative advantages over the ant colony optimization (ACO) algorithm.

The adaptive and reliable routing protocol called ARdeep is proposed in [142]. This deep learning-based protocol autonomously distinguishes network variations using the MDP model. In the proposed model, a node holding a packet determines its state and takes action to find the next-hop node using DQN. Factors such as Packet Error Rate, link status, connection time, and nodes' remaining energy influence routing decisions. For the packet delivery ratio and the end-to-end delay, the proposed ARdeep performs better than the Q-learning-based geographical protocol. However, the complexity of the proposed model is not discussed in the study.

The study by [143] introduces the Q-Learning-based Fuzzy Logic for Multi-Objective Routing Algorithm in Flying ad hoc Networks (QLFLMOR). QLFLMOR uses Q-learning and fuzzy logic in UAVs to select the optimal routing path based on link and path-level parameters. By including both link-level and path-level parameters, the algorithm provides a well-rounded approach to routing, balancing immediate link quality with overall path efficiency. Experimental results demonstrate that QLFLMOR achieves lower hop count and energy consumption compared to other routing algorithms. However, the integration of fuzzy logic and Q-learning adds to the complexity of the algorithm, which is computationally intensive for real-time applications.

To summarize, most routing protocols assume that the UAV networks are fully connected, but this is not the case in reality, and the broken links cause failure of routing protocols. In the conventional routing protocols, the node mobility is designed for 2D spaces, whereas the UAV moves in 3D space. In most studies, UAV mobility is converted

into 2D scenarios. The conventional Q-learning-based RL algorithm is used in almost all studies, and it causes average overhead maximization compared to conventional routing protocols. In the future, the RL-based method should be modified and implemented to accommodate nonlinear UAV movement for reliable transmission links. The research work in the area of AI-based UAV network routing for resource management is summarized in Table 14 based on the objectives of the research, the algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

Table 14. AI/ML-based routing solution for UAV networks.

Reference	Objective	Algorithm	Metric (s)
[136]	Group regions into clusters and find the best route	DRL	Optimality gap and temporal gap
[137]	UAV location optimization and relay path planning	RGNN	data rate achieved and time complexity.
[138]	Enhance end-to-end security through data encryption	Ant behavior	PDR, throughput and bandwidth utilization
[139]	Collision-free routing policies for UAVs	MARL	Average distance travelled and trajectories
[140]	Achieve lower latency by predicting future node positions	RL (PARRoT)	PDR
[141]	To identify adjacent nodes in real-time	FLRLR	Number of hops and link connectivity
[142]	Next hop selection	DQN based ARdeep	PDR and end-to-end delay
[143]	Select the optimal routing path based on link and path-level parameters	Q-learning based QLFLMOR	Hop count and energy consumption

5.4. AI for Routing in UAV-IoV Networks

Efficient data dissemination among vehicles and optimization of multi-hop path and relay selection is a complex task in IoV. The network latency and reliability considering the increasing vehicle density in future networks is crucial in making routing decisions. In this regard, the UAV-based routing in IoV is a relatively new area and has not been properly explored yet. The summary of AI/ML UAV-IoV routing protocols is provided in Table 15.

Table 15. AI/ML solution for UAV-IoV-based routing.

Reference	Learning Mechanism	Contribution	Evaluation
[144]	A Q-learning based load balancing routing (Q-LBR)	Q-Learning	Improved PDR, network utilization, and latency by more than 8%, 28%, and 30%.
[145]	UAV-assisted QAGR algorithm	Simulated in NS-3, Q-Learning	90% PDR achieved
[146]	Relay selection for A2G VANETs	Q-Learning	96% PDR achieved

In [144], the traffic congestion problem is dealt with by using a Q-learning-based load balancing routing (Q-LBR). It estimated the network load using a low-overhead technique to estimate the network load through the queue status of ground vehicular nodes and Q-learning based load balancing based on the current traffic condition. It finally implements a reward control function for Q-learning convergence by considering the UAV relay node's load and ground network congestion. The simulation results show that Q-LBR achieves better PDR, network utilization, and latency compared to the traditional routing protocols. Overall, the paper is well structured and provides a comprehensive analysis of the proposed method, supported by extensive simulation results.

In [145], an adaptive UAV-assisted geographic routing with Q-Learning (QAGR) is proposed. Routing is performed using two different components namely aerial and base components. UAVs use the combination of the fuzzy-logic and depth-first-search (DFS) algorithms to find the global routing path. This routing path information is transferred to the requesting vehicle on the ground. A fixed-sized Q-table is maintained at the vehicle, which is updated with the global routing path. The proposed QAGR routing protocol is evaluated using end-to-end delay, packet delivery ratio, and hop count as metrics. The end-to-end delay of the QAGR is the highest among all the comparative routing protocols. This clearly indicates that the convergence of the proposed algorithm is very slow, and it is not discussed in the study.

In [146], the authors address the relay selection problem for UAV-based VANET. They formulated the relay selection problem involving the state transition probabilities and transmission consumption (STP-TC) trade-off as a multi-objective optimization problem. The STP and TC are modeled from the source node to the destination node. Next, the STP threshold is set up. Finally, the Q-learning technique is employed to solve the proposed multi-objective optimization problem. This study is unique from all the other studies discussed so far as it considers various UAV heights and their impact on latency and delivery ratio. The proposed protocol outperforms all the other routing protocols. Moreover, the authors change the complexity of the protocol by increasing the number of vehicles, and the proposed STPTC protocol gives satisfactory performance.

In UAV-based IoV routing protocol research, control messages are periodically exchanged, and the flooding of routing messages leads to excessive bandwidth consumption and high overhead. Moreover, most routing protocols developed using the RL model prioritize QoS requirements. Researchers should consider incorporating additional objectives like link quality and delay into the reward function to ensure rapid model convergence and smoother operation. The research work in the area of AI-based UAV-assisted IoV network routing for resource management is summarized in Table 15 based on the objectives of the research, algorithm designed, and the metrics used to evaluate the performance of the proposed algorithm.

This section addresses the constraints of ML/AI algorithms and the simulation software utilized for training these algorithms.

6. Major Limitations and Challenges in AI/ML Deployment

This section addresses the constraints of ML/AI algorithms and the simulation software utilized for training these algorithms. From the detailed review and critical discussions in Sections 4 and 5, including Sections 4.1, 4.2, and 4.3 and, Sections 5.2, 5.3, and 5.4 respectively, it is evident that in last five years, numerous AI/ML-based IoV resource management and routing algorithms have been proposed and implemented to improve the performance of the UAV and IoV networks. While AI/ML approaches are data-driven and can yield fairly accurate solutions in most cases, they also have several limitations. Key limitations of ML and DL include

- **Application Specificity:** ML models are tailored to specific applications. For instance, a DL model trained on vehicular applications like network congestion prediction [147] or classification [148] will perform well in that domain but may not effectively predict or classify traffic congestion in a different contexts.
- **Noisy and Incomplete Data:** ML agents often encounter noisy and incomplete data [77,78], adversely affecting their learning and decision-making capabilities.
- **Explainability:** Interpreting and explaining the decisions made by ML can be challenging, particularly when they control physical real world systems that can have real-world consequences.

Moreover, the ML and DL models rely heavily on data, and their effectiveness is contingent on data availability. Most DL algorithms require substantial data. However, in UAV-based vehicular communications, historical data for time-sensitive tasks such as resource management, mobility prediction, and routing decisions are often scarce. Thus,

there is an imperative need for open-source and reliable data pertaining to UAV vehicles, including mechanisms to produce and estimate the accurate dataset size needed to train and test ML and DL algorithms.

This necessity has led most research to employ RL for task-offloading and routing tasks. RL and its variants have been proven to handle non-convex problems effectively, such as task management, energy efficiency, and routing. In RL, however, the agent's actions are contingent upon the received rewards or penalties. Specifically, in routing-related problems, as the state and action sets become large, it affects the convergence, increases the latency, and increases the dimensionality of the Q tables, which results in high memory consumption.

Recently, FL has become a trusted solution as it ensures data privacy and reduces time complexity. However, FL is vulnerable to backdoor attacks that can compromise the model's integrity by injecting poisoned data or models. Additionally, the convergence of the FL model presents another challenge, as it is specific to problems such as the convexity of the loss function and the frequency of model updates. Without adequate data, the model may not yield accurate results. Furthermore, the UAV-IoV network is diverse, comprising drones of various sizes and specifications, and vehicles with dynamic computational and processing capabilities, including different GPUs. Implementing FL in such a diverse network means that drones and vehicles will exhibit varying response times. During FL, model updates occur at each communication round, and any delays can lead to slow model convergence.

In summary, despite some limitations, AI/ML, including ML, RL, and FL-based solutions, have shown improved outcomes for resource management and routing in UAV-based IoV networks compared to other methods addressing non-convex vehicular and UAV network challenges.

7. Conclusions

UAV-based aerial networks introduce a third (spatial) dimension to wireless networks, particularly for IoVs. UAVs are distinctive compared to other static communication networks as they can function as mobile base stations, and their integration has rendered the network more dynamic. Their mobility introduces both versatility and complexity to vehicular networks. Consequently, traditional methods are inadequate, and AI/ML plays a crucial role in UAV-based IoV.

This paper offers a comprehensive comparative analysis of AI algorithms' applications within UAV-based IoV paradigms. We have examined various challenges, including resource management and routing techniques in UAV-based IoV, employing different AI strategies. AI-based algorithms have improved system performance over traditional methods. Notably, combining multiple AI algorithms to leverage their strengths yields nearly optimal solutions for resource management and routing in UAV-IoV networks. This also results in increased system throughput, reduced energy consumption, and decreased latency. Nonetheless, the significant computational resources required by AI algorithms in dynamic vehicular and UAV environments pose a substantial challenge. In a conventional cellular system, when vehicles move from one communication cell to another, a handover takes place between base stations. The vehicles are fast-moving, and the connection time is very short, which results in lower data rates and overheads. Emerging MEC and VEC architectures often offload computing tasks to RSUs or Fog nodes; however, this offloading also suffers from high latency based on the size of the computational task to be offloaded to the UAV acting as the MEC. In this regard, we intend to include UAV-based access points for IoV communication in Cell-Free massive Multiple-Input and Multiple-Output (CF-mMIMO) contexts. In cell-free communication, more than one UAV serving as the aerial access point can serve a vehicle at the same time, and this significantly improves the achievable data rate at very low latency. We aim to introduce AI into CF-mMIMO to achieve better computational offloading between vehicles and UAVs so that our model will be capable of autonomous operation, enhanced connectivity, and robustness.

Author Contributions: Conceptualization, S.A.A.S. and X.F.; data curation, S.A.A.S.; formal analysis, S.A.A.S.; funding acquisition, X.F. and R.K.; investigation, S.A.A.S.; project administration, X.F. and R.K.; supervision, X.F. and R.K.; validation, X.F. and R.K.; visualization, S.A.A.S.; writing—original draft, S.A.A.S.; writing—review and editing, S.A.A.S. and X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Natural Sciences and Engineering Research Council (NSERC) and Toronto Metropolitan University Canada.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

3DQN	Double Dueling Deep Q Network	LTE	Long Term Evolution
3GPP	3rd Generation Partnership Program	MAC	Medium Access Control
5G	5th Generation	MARL	Multi-Agent Reinforcement Learning
6G	6th Generation	MACEL	Multi-Agent Collaborative Environment Learning
A3C	Asynchronous Actor Critic	MA2DDPG	Multi-Attentive Deep Deterministic Policy Gradient
AI	Auction Integration	MADQL	Multi Agent Deep Q-Learning
AI	Artificial Intelligence	MADRL	Multi Agent Deep Reinforcement Learning
AC-Mix	actor-critic Mixing Network	MDP	Markov Decision Process
AODV	Ad hoc On-Demand Distance Vector	MEC	Mobile Edge Computing
AEC	Average Energy Constraint	MINLP	Mixed integer non linear programming
ACO	Ant Colony Optimization	MJDDPG	Multi-objective Joint Optimization-Oriented DDPG Algorithm
ASR-CUMS	Automated Slice Resource Control and Update Management System	MLHR	Multi Hierarchical Routing
ARdeep	Adaptive Reliable Deep	MLP	Multi-Layer Perceptron
BT-MP-DQN	Beam-forming Control and Trajectory-Multi-Pass Deep Q Network	ML	Machine Learning
BS	Base Station	mmWAVE	Millimeter Wave
CA-MOEA	Clustering based Adoptive Multi Objective Evolutionary Algorithm	MOEA/D	A Multi-objective Evolutionary Algorithm Based on Decomposition
CCSRL	Cluster-enabled Cooperative Scheduling based on Reinforcement Learning	MSA-LS	Mobile Service Amount based Link scheduling
CEPF	Context Aware Packet Forwarding	NGSIM	Next generation Simulation
CF-mMIMO	Cell Free Massive Multiple Input Multiple output	NP	Non-Deterministic Polynomials
CH	Cluster Head	NOMA	Non-Orthogonal Multiple Access
CKF	Constant Kalman Filter	OFDMA	Orthogonal Frequency-Division Multiple Access
CNN	Convolution Neural Networks	OLSR	Optimized Linked State Routing
		OMNET++	Objective Modular network Testbed in C++
CSMA	Carrier Sense Multiple Access	PARROT	Predictive ad hoc Routing Combined with Reinforcement Learning and Trajectory Knowledge
CSS	Cooperative Spectrum Sensing	PDR	Packet Delivery Ratio
D3QN	Deep Double Duelling Q Network	PPO	Proximal Policy Optimization
DCR	Data Centring Routing	QAGR	Geographic Routing with Q-Learning
DDPG	Deep Deterministic Policy Gradient	QFHR	Q-learning and Fuzzy-based Hierarchical Routing Solution
DFMDP	Discrete Time and Finite-State Markov Decision Process	QLBR	Q-Learning based Load Balancing Routing
DFS	Depth First Search	QOE	Quality of Experience
DGCIM	Dual Graph Coloring based Interference Management	QLFMOR	Q-Learning based Fuzzy Logic for Multi Objective Routing Algorithm
		QTAR	Q-Learning based Traffic Aware Routing

DNN	Deep Neural Network	QoS	Quality of Service
DL	Deep Learning	RF	Random Forest
DQL	Deep Q-Learning	RGNN	Reinforcement Graph Neural Network
DQN	Deep Q-Network	RIS	Re-configurable Intelligent Surface
DRL-RASO	Deep Reinforcement Learning based Resource Allocation and Speed Optimization	RL	Reinforcement Learning
DRL	Deep Reinforcement Learning	RRPV	Reinforcement Learning Routing Protocol for Vehicles
DSRC	Dedicated Short Range Communications	SGD	Stochastic Gradient Descent
DRQN	Deep Recurrent Object Networks	SI	Swarm Intelligence
DSDV	Destination Sequenced Distance Vector	SNR	Signal to Noise Ratio
DSR	Dynamic Source Routing	STPTC	State Transition Probabilities and Transmission Consumption
DTN	Delay Tolerant Networking	SUMO	Simulation of Urban Mobility
EED	End-to-End Delay	SU	Secondary User
FANET	Flying ad hoc Network	SVM	Support Vector Machine
FL	Federated Learning	SWIPT	Simultaneous Wireless and power Transfer
FLRLR	Fuzzy Logic Reinforcement Learning based Routing	TD3	Twin Delayed Deep Deterministic Policy Gradient
GCS	Ground Control Station	UAS	Unmanned Aerial System
GMM	Gaussian Mixture Model	UAV	Unmanned Aerial Vehicle
GRNN	Generalized Regression Neural Network	UE	User Equipment
GPS	Global Positioning System	UCPA	UAV based Clustering and Positioning Protocol
GNNRL	Graph Neural Network based on Reinforcement Learning	URLLC	Ultra-Reliable Low-Latency Communications
GPGC-RLF	Grouping Graph Coloring with Recursive Largest First	V2I	Vehicle-to-Infrastructure
GPSR	Greedy Perimeter Stateless Routing	V2N	Vehicle-to-Network
GS	Grid Search	V2P	Vehicle-to-Pedestrian
GYGC	Greedy Graph Coloring	V2R	Vehicle-to-Roadside Infrastructure
HAB	High-Altitude Balloon	V2V	Vehicle-to-Vehicle
HAP	High Altitude Platform	V2X	Vehicle-to-Everything
IMU	Inertial Measurement Unit	V2U	Vehicle-to-Unmanned Aerial Vehicle
IoD	Internet of Drones	VANET	Vehicular ad hoc Network
IoT	Internet of Things	VEC	Vehicular Edge Computing
iProPHET	Improved Probability Routing Protocol using History of Encounters and Transitivity	VFC	Vehicular Fog Clouds
IQR	Intersection- based Q-Learning	VUE	Vehicle User Equipment
IQS	Improved Q-learning	VM	Virtual Machines
ITS	Intelligent Transport Systems	VFRM	Vehicular Fog Resource Management
JTSM	Joint Time Series Modeling	VNF	Virtual Network Functions
KKT	Karush–Kuhn–Tucker	WPT	Wireless Power Transfer
LBTO	Load Balancing and Task Offloading	WMMSE	Weighted Minimum Mean Square Error
LCDR	Load Carrying and Delivery Routing	WSN	Wireless Sensor Network
LIDAR	Light Detection and Ranging	ZRP	Zone Routing Protocol
LOS	Line-of-Sight		
LPA	Long Prediction Algorithm		
LSTM	Long Short Term Memory		

References

1. Hashemi, S.; Zarei, M. Internet of Things backdoors: Resource management issues, security challenges, and detection methods. *Trans. Emerg. Telecommun. Technol.* **2021**, *32*, e4142. [CrossRef]
2. Alexander, G. What is Internet of things (IoT)? *IOT Agenda* 2021. Available online: <https://www.rtsrl.eu/blog/what-is-internet-of-things-iot/> (accessed on 3 August 2023).
3. Tang, C.; Wei, X.; Liu, C.; Jiang, H.; Wu, H.; Li, Q. UAV-Enabled Social Internet of Vehicles: Roles, Security Issues and Use Cases. In *Security and Privacy in Social Networks and Big Data. SocialSec 2020. Communications in Computer and Information Science*; Xiang, Y., Liu, Z., Li, J., Eds.; Springer: Singapore, 2020; Volume 1298.

4. Jamalzadeh, M.; Maadani, M.; Mahdavi, M. EC-MOPSO: An edge computing-assisted hybrid cluster and MOPSO-based routing protocol for the Internet of Vehicles. *Ann. Telecommun.* **2022**, *77*, 491–503. [\[CrossRef\]](#)
5. Krishna, M. A Survey UAV-Assisted VANET Routing Protocol. *Int. J. Comput. Sci. Trends Technol.* **2020**, *8*, 68–74.
6. Guerna, A.; Bitam, S.; Calafate, C.T. Roadside Unit Deployment in Internet of Vehicles Systems: A Survey. *Sensors* **2022**, *22*, 3190. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Ghazal, T.M.; Hasan, M.K.; Alshurideh, M.T.; Alzoubi, H.M.; Ahmad, M.; Akbar, S.S.; Al Kurdi, B.; Akour, I.A. IoT for Smart Cities: Machine Learning Approaches in Smart Healthcare—A Review. *Future Internet* **2021**, *13*, 218. [\[CrossRef\]](#)
8. Yaqoob, S.; Ullah, A.; Awais, M.; Katib, I.; Albeshri, A.; Mehmood, R.; Rodrigues, J.J. Novel congestion avoidance scheme for Internet of Drones. *Computer Commun.* **2021**, *169*, 202–210. [\[CrossRef\]](#)
9. Mazyavkina, N.; Sviridov, S.; Ivanov, S.; Burnaev, E. Reinforcement learning for combinatorial optimization: A survey. *Comput. Oper. Res.* **2021**, *134*, 105400. [\[CrossRef\]](#)
10. Saravanan, M.; Ganeshkumar, P. Routing using reinforcement learning in vehicular ad hoc networks. *Comput. Intell.* **2020**, *36*, 682–697. [\[CrossRef\]](#)
11. Sun, Y.; Lin, Y.; Tang, Y. A Reinforcement Learning-Based Routing Protocol in VANETs. In *Communications, Signal Processing, and Systems; CSPS 2017; Lecture Notes in Electrical Engineering*; Liang, Q., Mu, J., Jia, M., Wang, W., Feng, X., Zhang, B., Eds.; Springer: Singapore, 2019; Volume 463.
12. Liang, L.; Ye, H.; Li, G.Y. Toward Intelligent Vehicular Networks: A Machine Learning Framework. *IEEE Internet Things J.* **2019**, *6*, 124–135. [\[CrossRef\]](#)
13. Tong, W.; Hussain, A.; Bo, W.X.; Maharjan, S. Artificial Intelligence for Vehicle-to-Everything: A Survey. *IEEE Access* **2019**, *7*, 10823–10843. [\[CrossRef\]](#)
14. Tang, F.; Kawamoto, Y.; Kato, N.; Liu, J. Future Intelligent and Secure Vehicular Network Toward 6G: Machine-Learning Approaches. *Proc. IEEE* **2020**, *108*, 292–307. [\[CrossRef\]](#)
15. Tang, F.; Mao, B.; Kato, N.; Gui, G. Comprehensive Survey on Machine Learning in Vehicular Network: Technology, Applications and Challenges. *IEEE Commun. Surv. Tutorials* **2021**, *23*, 2027–2057. [\[CrossRef\]](#)
16. Hossain, M.A.; Noor, R.M.; Yau, K.L.A.; Azzuhri, S.R.; Z'aba, M.R.; Ahmedy, I. Comprehensive Survey of Machine Learning Approaches in Cognitive Radio-Based Vehicular Ad Hoc Networks. *IEEE Access* **2020**, *8*, 78054–78108. [\[CrossRef\]](#)
17. Du, Z.; Wu, C.; Yoshinaga, T.; Yau, K.L.A.; Ji, Y.; Li, J. Federated Learning for Vehicular Internet of Things: Recent Advances and Open Issues. *IEEE Open J. Comput. Soc.* **2020**, *1*, 45–61. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Ali, E.S.; Hasan, M.K.; Hassan, R.; Saeed, R.A.; Hassan, M.B.; Islam, S.; Bevinakoppa, S. Machine Learning Technologies for Secure Vehicular Communication in Internet of Vehicles: Recent Advances and Applications. *Secur. Commun. Netw.* **2021**, *2021*, 8868355. [\[CrossRef\]](#)
19. Nurcahyani, I.; Lee, J.W. Role of Machine Learning in Resource Allocation Strategy over Vehicular Networks: A Survey. *Sensors* **2021**, *21*, 6542. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Mekrache, A.; Bradai, A.; Moulay, E.; Dawaliby, S. Deep reinforcement learning techniques for vehicular networks: Recent advances and future trends towards 6G. *Veh. Commun.* **2022**, *33*, 100398. [\[CrossRef\]](#)
21. Gillani, M.; Niaz, H.A.; Tayyab, M. Role of Machine Learning in WSN and VANETs. *Int. J. Electr. Comput. Eng. Res.* **2021**, *1*, 15–20. [\[CrossRef\]](#)
22. Mchergui, A.; Moulahi, T.; Zeadally, S. Survey on Artificial Intelligence (AI) techniques for Vehicular ad hoc Networks (VANETs). *Veh. Commun.* **2022**, *34*, 100403. [\[CrossRef\]](#)
23. Noor-A-Rahim, M.; Liu, Z.; Lee, H.; Ali, G.M.N.; Pesch, D.; Xiao, P. A Survey on Resource Allocation in Vehicular Networks. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 701–721. [\[CrossRef\]](#)
24. Lansky, J.; Rahmani, A.M.; Hosseinzadeh, M. Reinforcement Learning-Based Routing Protocols in Vehicular Ad Hoc Networks for Intelligent Transport System (ITS): A Survey. *Mathematics* **2022**, *10*, 4673. [\[CrossRef\]](#)
25. Javed, A.R.; Hassan, M.A.; Shahzad, F.; Ahmed, W.; Singh, S.; Baker, T.; Gadekallu, T.R. Integration of Blockchain Technology and Federated Learning in Vehicular (IoT) Networks: A Comprehensive Survey. *Sensors* **2022**, *22*, 4394. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Christopoulou, M.; Barmounakis, S.; Koumaras, H.; Kaloxylou, A. Artificial Intelligence and Machine Learning as key enablers for V2X communications: A comprehensive survey. *Veh. Commun.* **2023**, *39*, 100569. [\[CrossRef\]](#)
27. Hasan, M.K.; Jahan, N.; Nazri, M.Z.A.; Islam, S.; Khan, M.A.; Alzahrani, A.I.; Nam, Y. Federated Learning for Computational Offloading and Resource Management of Vehicular Edge Computing in 6G-V2X Network. *IEEE Trans. Consum. Electron.* **2024**, *70*, 3827–3847. [\[CrossRef\]](#)
28. Hemmati, A.; Zarei, M.; Souri, A. UAV-based Internet of Vehicles: A systematic literature review. *Intell. Syst. Appl.* **2023**, *18*, 200226. [\[CrossRef\]](#)
29. Heidari, A.; Jafari Navimipour, N.; Unal, M.; Zhang, G. Machine Learning Applications in Internet-of-Drones: Systematic Review, Recent Deployments, and Open Issues. *ACM Comput. Surv.* **2023**, *55*, 1–45. [\[CrossRef\]](#)
30. Sun, C.; Fontanesi, G.; Canberk, B.; Mohajerzadeh, A.; Chatzinotas, S.; Grace, D.; Ahmadi, H. Advancing UAV Communications: A Comprehensive Survey of Cutting-Edge Machine Learning Techniques. *IEEE Open J. Veh. Technol.* **2024**, *5*, 825–854. [\[CrossRef\]](#)
31. Banafaa, M.; Pepeoglu, Ö.; Shaye, I.; Alhammedi, A.; Shamsan, Z.; Razaz, M.A.; Al-Sowayan, S. A comprehensive survey on 5G-and-beyond networks with UAVs: Applications, emerging technologies, regulatory aspects, research trends and challenges. *IEEE Access* **2024**, *12*, 7786–7826. [\[CrossRef\]](#)

32. Sharma, S.; Kaushik, B. A survey on Internet of vehicles: Applications, security issues and solutions. *Veh. Commun.* **2019**, *20*, 100182. [\[CrossRef\]](#)
33. Chaurasia, R.; Mohindru, V. Unmanned Aerial Vehicle (UAV): A comprehensive survey. In *Unmanned Aerial Vehicles for Internet of Things (IoT): Concepts, Techniques, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2021; pp. 1–27. [\[CrossRef\]](#)
34. IEEE P1609.0/D9; IEEE Draft Guide for Wireless Access in Vehicular Environments (WAVE)—Architecture. IEEE: Piscataway, NJ, USA, 2017.
35. 802.11u-2011; IEEE Standard for Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Networks-Specific Requirements-Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Amendment 9: Interworking with External Networks; In Amendment to IEEE Std 802.11-2007. IEEE: Piscataway, NJ, USA, 2011; pp. 1–208. [\[CrossRef\]](#)
36. Li, J.; Shi, M.; Li, J.; Yao, D. Media Access Process Modeling of LTE-V-Direct Communication Based on Markov Chain. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018; pp. 61–66. [\[CrossRef\]](#)
37. Hassan, N.; Fernando, X.; Woungang, I. An Emergency Message Routing Protocol for Improved Congestion Management in Hybrid RF/VLC VANETs. *Telecom* **2024**, *5*, 21–47. [\[CrossRef\]](#)
38. Khan, L.U. Visible light communication: Applications, architecture, standardization and research challenges. *Digit. Commun. Netw.* **2017**, *3*, 78–88. [\[CrossRef\]](#)
39. Cen, N.; Jagannath, J.; Moretti, S.; Guan, Z.; Melodia, T. LANET:Visible-light ad hoc networks. *Ad Hoc Netw.* **2019**, *84*, 107–123. [\[CrossRef\]](#)
40. Fernando, X.; Hasan, F. *Visible Light Communications—Vehicular Applications*; IOP Publishing Ltd.: Bristol, UK, 2019; ISBN 978-0-7503-2284-3.
41. Obaid, A.; Fernando, X.; Jaseemuddin, M. A mobility-aware cluster-based MAC protocol for radio-frequency energy harvesting cognitive wireless sensor networks. *IET Wirel. Sens. Syst.* **2021**, *11*, 206–218. [\[CrossRef\]](#)
42. Choi, J.; Va, V.; Gonzalez-Prelcic, N.; Daniels, R.; Bhat, C.R.; Heath, R.W. Millimeter-wave vehicular communication to support massive automotive sensing. *IEEE Commun. Mag.* **2016**, *54*, 160–167. [\[CrossRef\]](#)
43. Va, V.; Shimizu, T.; Bansal, G.; Heath, R.W., Jr. *Millimeter Wave Vehicular Communications: A Survey*; Now: Hanover, MA, USA, 2016.
44. Araniti, G.; Campolo, C.; Condoluci, M.; Iera, A.; Molinaro, A. LTE for vehicular networking: A survey. *IEEE Commun. Mag.* **2013**, *51*, 148–157. [\[CrossRef\]](#)
45. Papathanassiou, A.; Khoryaev, A. Cellular V2X as the essential enabler of superior global connected transportation services. *IEEE 5G Tech. Focus* **2017**, *1*, 1–2.
46. PC5. Initial Cellular V2X Standard Completed. 2018. Available online: <https://www.3gpp.org/news-events/3gpp-news/v2x-r14> (accessed on 8 August 2023).
47. Husain, S.; Kunz, A.; Prasad, A.; Pateromichelakis, E.; Samdanis, K.; Song, J. The Road to 5G V2X: Ultra-High Reliable Communications. In Proceedings of the IEEE Conference on Standards for Communications and Networking (CSCN), Paris, France, 29–31 October 2018; pp. 1–6. [\[CrossRef\]](#)
48. Osorio, D.P.M.; Ahmad, I.; Sánchez, J.D.V.; Gurtov, A.; Scholliers, J.; Kuttila, M.; Porambage, P. Towards 6G-Enabled Internet of Vehicles: Security and Privacy. *IEEE Open J. Commun. Soc.* **2022**, *3*, 82–105. [\[CrossRef\]](#)
49. Commission Delegated Regulation (EU) 2019/945, 2019, Official Journal of the European Union, 12 March 2019. Available online: https://eur-lex.europa.eu/eli/reg_del/2019/945/oj (accessed on 10 December 2023).
50. Altawy, R.; Youssef, A.M. Security, privacy, and safety aspects of civilian drones: A survey. *ACM Trans.-Cyber-Phys. Syst.* **2016**, *1*, 1–25. [\[CrossRef\]](#)
51. Villa, T.F.; Salimi, F.; Morton, K.; Morawska, L.; Gonzalez, F. Development and Validation of a UAV Based System for Air Pollution Measurements. *Sensors* **2016**, *16*, 2202. [\[CrossRef\]](#)
52. Chao, H.; Cao, Y.; Chen, Y. Autopilots for small unmanned aerial vehicles: A survey. *Int. J. Control. Autom. Syst.* **2010**, *8*, 36–44. [\[CrossRef\]](#)
53. Höflinger, F.; Müller, J.; Zhang, R.; Reindl, L.M.; Burgard, W. A Wireless Micro Inertial Measurement Unit (IMU). *IEEE Trans. Instrum. Meas.* **2013**, *62*, 2583–2595. [\[CrossRef\]](#)
54. Vasylenko, M.P. Telemetry System of Unmanned Aerial Vehicles. *Electron. Control. Syst.* **2018**, *3*, 95–100. [\[CrossRef\]](#)
55. Liu, Y.; Dai, H.N.; Wang, Q.; Shukla, M.K.; Imran, M. Unmanned aerial vehicle for Internet of everything: Opportunities and challenges. *Comput. Commun.* **2020**, *155*, 66–83. [\[CrossRef\]](#)
56. Chriki, A.; Touati, H.; Snoussi, H.; Kamoun, F. FANET: Communication, mobility models and security issues. *Comput. Netw.* **2019**, *163*, 106877. [\[CrossRef\]](#)
57. Ad Hoc Network, NIST. Available online: <https://csrc.nist.gov/glossary> (accessed on 15 December 2023).
58. Xia, Y.; Wu, L.; Wang, Z.; Zheng, X.; Jin, J. Cluster-Enabled Cooperative Scheduling Based on Reinforcement Learning for High-Mobility Vehicular Networks. *IEEE Trans. Veh. Technol.* **2020**, *69*, 12664–12678. [\[CrossRef\]](#)
59. Zhang, X.; Peng, M.; Yan, S.; Sun, Y. Deep-Reinforcement-Learning-Based Mode Selection and Resource Allocation for Cellular V2X Communications. *IEEE Internet Things J.* **2020**, *7*, 6380–6391. [\[CrossRef\]](#)
60. Khan, Z.; Fan, P.; Abbas, F.; Chen, H.; Fang, S. Two-Level Cluster Based Routing Scheme for 5G V2X Communication. *IEEE Access* **2019**, *7*, 16194–16205. [\[CrossRef\]](#)

61. Li, F.; Song, X.; Chen, H.; Li, X.; Wang, Y. Hierarchical Routing for Vehicular Ad Hoc Networks via Reinforcement Learning. *IEEE Trans. Veh. Technol.* **2019**, *68*, 1852–1865. [\[CrossRef\]](#)
62. Liang, L.; Ye, H.; Li, G.Y. Spectrum Sharing in Vehicular Networks Based on Multi-Agent Reinforcement Learning. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 2282–2292. [\[CrossRef\]](#)
63. Alatabani, L.E.; Saeed, R.A.; Ali, E.S.; Mokhtar, R.A.; Khalifa, O.O.; Hayder, G. Vehicular network spectrum allocation using hybrid NOMA and multi-agent reinforcement learning. In *Sustainability Challenges and Delivering Practical Engineering Solutions: Resources, Materials, Energy, and Buildings*; Springer International Publishing: Cham, Switzerland, 2023; pp. 151–158.
64. Paul, A.; Choi, K. Deep learning-based selective spectrum sensing and allocation in cognitive vehicular radio networks. *Veh. Commun.* **2023**, *41*, 100606. [\[CrossRef\]](#)
65. Pan, Q.; Wu, J.; Nebhen, J.; Bashir, A.K.; Su, Y.; Li, J. Artificial intelligence-based energy efficient communication system for intelligent reflecting surface-driven vanets. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19714–19726. [\[CrossRef\]](#)
66. Xu, Y.-H.; Yang, C.-C.; Hua, M.; Zhou, W. Deep Deterministic Policy Gradient (DDPG)-Based Resource Allocation Scheme for NOMA Vehicular Communications. *IEEE Access* **2020**, *8*, 18797–18807. [\[CrossRef\]](#)
67. Ye, H.; Li, G.Y.; Juang, B.H.F. Deep Reinforcement Learning Based Resource Allocation for V2V Communications. *IEEE Trans. Veh. Technol.* **2019**, *68*, 3163–3173. [\[CrossRef\]](#)
68. Wang, Y.; Li, X.; Wan, P.; Shao, R. Intelligent dynamic spectrum access using deep reinforcement learning for VANETs. *IEEE Sens. J.* **2021**, *21*, 15554–15563. [\[CrossRef\]](#)
69. Kumar, A.S.; Zhao, L.; Fernando, X. Multi-Agent Deep Reinforcement Learning-Empowered Channel Allocation in Vehicular Networks. *IEEE Trans. Veh. Technol.* **2022**, *71*, 1726–1736. [\[CrossRef\]](#)
70. Kumar, A.S.; Zhao, L.; Fernando, X. Mobility Aware Channel Allocation for 5G Vehicular Networks using Multi-Agent Reinforcement Learning. In Proceedings of the ICC 2021—IEEE International Conference on Communications, Montreal, QC, Canada, 14–18 June 2021; pp. 1–6. [\[CrossRef\]](#)
71. Xing, Y.; Lv, C.; Cao, D. Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles. *IEEE Trans. Veh. Technol.* **2019**, *69*, 1341–1352. [\[CrossRef\]](#)
72. Hou, L.; Lei, L.; Zheng, K.; Wang, X. AQ-Learning-Based Proactive Caching Strategy for Non-Safety Related Services in Vehicular Networks. *IEEE Internet Things J.* **2019**, *6*, 4512–4520. [\[CrossRef\]](#)
73. Ding, W.; Shen, S. Online Vehicle Trajectory Prediction using Policy Anticipation Network and optimization-based Context Reasoning. In Proceedings of the International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 9610–9616. [\[CrossRef\]](#)
74. Dai, S.; Li, L.; Li, Z. Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction. *IEEE Access* **2019**, *7*, 38287–38296. [\[CrossRef\]](#)
75. Chandra, R.; Bhattacharya, U.; Bera, A.; Manocha, D. Taphic: Trajectory prediction in dense and heterogeneous traffic using weighted interactions. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8483–8492.
76. Xie, G.; Shangguan, A.; Fei, R.; Ji, W.; Hei, X. Motion trajectory prediction based on a CNN-LSTM sequential model. *Sci. China Inf. Sci.* **2020**, *63*, 1–21. [\[CrossRef\]](#)
77. Cui, Y.; Liang, Y.; Wang, R. Resource Allocation Algorithm With Multi-Platform Intelligent Offloading in D2D-Enabled Vehicular Networks. *IEEE Access* **2019**, *7*, 21246–21253. [\[CrossRef\]](#)
78. Saleh, A.H.; Anpalagan, A. AI Empowered Computing Resource Allocation in Vehicular ad hoc NETWORKS. In Proceedings of the 2022 7th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, 19–20 May 2022; pp. 221–226.
79. Lee, S.-S.; Lee, S. Resource Allocation for Vehicular Fog Computing Using Reinforcement Learning Combined with Heuristic Information. *IEEE Internet Things J.* **2020**, *7*, 10450–10464. [\[CrossRef\]](#)
80. Haris, M.; Shah, M.A.; Maple, C. Internet of intelligent vehicles (IoIV): An intelligent VANET based computing via predictive modeling. *IEEE Access* **2023**, *11*, 49665–49674. [\[CrossRef\]](#)
81. Ibrar, M.; Akbar, A.; Jan, S.R.U.; Jan, M.A.; Wang, L.; Song, H.; Shah, N. Artnet: Ai-based resource allocation and task offloading in a reconfigurable Internet of vehicular networks. *IEEE Trans. Netw. Sci. Eng.* **2020**, *9*, 67–77. [\[CrossRef\]](#)
82. Tayyaba, S.K.; Khattak, H.A.; Almogren, A.; Shah, M.A.; Din, I.U.; Alkhalifa, I.; Guizani, M. 5G Vehicular Network Resource Management for Improving Radio Access Through Machine Learning. *IEEE Access* **2020**, *8*, 6792–6800. [\[CrossRef\]](#)
83. Muhammad, A.; Khan, T.A.; Abbass, K.; Song, W.-C. An End-to-end Intelligent Network Resource Allocation in IoV: A Machine Learning Approach. In Proceedings of the 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall), Victoria, BC, Canada, 4–7 October 2020; pp. 1–5. [\[CrossRef\]](#)
84. Zhu, X.; Luo, Y.; Liu, A.; Bhuiyan, M.Z.A.; Zhang, S. Multiagent deep reinforcement learning for vehicular computation offloading in iot. *IEEE Internet Things J.* **2021**, *8*, 9763–9773. [\[CrossRef\]](#)
85. Kumar, A.S.; Zhao, L.; Fernando, X. Task Offloading and Resource Allocation in Vehicular Networks: A Lyapunov-Based Deep Reinforcement Learning Approach. *IEEE Trans. Veh. Technol.* **2023**, *72*, 13360–13373. [\[CrossRef\]](#)
86. Dai, Z.; Zhang, Y.; Zhang, W.; Luo, X.; He, Z. A Multi-Agent Collaborative Environment Learning Method for UAV Deployment and Resource Allocation. *IEEE Trans. Signal Inf. Process. Over Netw.* **2022**, *8*, 120–130. [\[CrossRef\]](#)

87. Alfaia, R.D.; Souto, A.V.d.F.; Cardoso, E.H.S.; Araújo, J.P.L.d.; Francês, C.R.L. Resource Management in 5G Networks Assisted by UAV Base Stations: Machine Learning for Overloaded Macrocell Prediction Based on Users' Temporal and Spatial Flow. *Drones* **2022**, *6*, 145. [\[CrossRef\]](#)
88. Khalili, A.; Monfared, E.M.; Zargari, S.; Javan, M.R.; Yamchi, N.M.; Jorswieck, E.A. Resource Management for Transmit Power Minimization in UAV-Assisted RIS HetNets Supported by Dual Connectivity. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 1806–1822. [\[CrossRef\]](#)
89. Lyu, T.; Zhang, H.; Xu, H. Resource Allocation in UAV-Assisted Wireless Powered Communication Networks for Urban Monitoring. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 7730456. [\[CrossRef\]](#)
90. Anicho, O.; Charlesworth, P.B.; Baicher, G.S.; Nagar, A.; Buckley, N. Comparative study for coordinating multiple unmanned HAPS for communications area coverage. In Proceedings of the 2019 International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, 11–14 June 2019; pp. 467–474.
91. Lin, Y.; Wang, M.; Zhou, X.; Ding, G.; Mao, S. Dynamic spectrum interaction of UAV flight formation communication with priority: A deep reinforcement learning approach. *IEEE Trans. Cogn. Commun. Netw.* **2020**, *6*, 892–903. [\[CrossRef\]](#)
92. Yang, C.; Liu, B.; Li, H.; Li, B.; Xie, K.; Xie, S. Learning Based Channel Allocation and Task Offloading in Temporary UAV-Assisted Vehicular Edge Computing Networks. *IEEE Trans. Veh. Technol.* **2022**, *71*, 9884–9895. [\[CrossRef\]](#)
93. Zeng, T.; Semiari, O.; Mozaffari, M.; Chen, M.; Saad, W.; Bennis, M. Federated Learning in the Sky: Joint Power Allocation and Scheduling with UAV Swarms. In Proceedings of the 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6. [\[CrossRef\]](#)
94. Liu, C.; Zhu, Q. Joint Resource Allocation and Learning Optimization for UAV-Assisted Federated Learning. *Appl. Sci.* **2023**, *13*, 3771. [\[CrossRef\]](#)
95. Deng, C.; Fang, X.; Wang, X. UAV-Enabled Mobile-Edge Computing for AI Applications: Joint Model Decision, Resource Allocation, and Trajectory Optimization. *IEEE Internet Things J.* **2023**, *10*, 5662–5675. [\[CrossRef\]](#)
96. Ji, P.; Jia, J.; Chen, J.; Guo, L.; Du, A.; Wang, X. Reinforcement learning based joint trajectory design and resource allocation for RIS-aided UAV multicast networks. *Comput. Netw.* **2023**, *227*, 109697. [\[CrossRef\]](#)
97. Li, Y.; Aghvami, A.H. Radio Resource Management for Cellular-Connected UAV: A Learning Approach. *IEEE Trans. Commun.* **2023**, *71*, 2784–2800. [\[CrossRef\]](#)
98. Munaye, Y.Y.; Juang, R.-T.; Lin, H.-P.; Tarekegn, G.B.; Lin, D.-B. Deep Reinforcement Learning Based Resource Management in UAV-Assisted IoT Networks. *App. Sci.* **2021**, *11*, 2163. [\[CrossRef\]](#)
99. Cui, J.; Liu, Y.; Nallanathan, A. Multi-agent reinforcement learning-based resource allocation for UAV networks. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 729–743. [\[CrossRef\]](#)
100. Zhu, S.; Gui, L.; Cheng, N.; Zhang, Q.; Sun, F.; Lang, X. UAV-enabled computation migration for complex missions: A reinforcement learning approach. *IET Commun.* **2020**, *14*, 2472–2480. [\[CrossRef\]](#)
101. Kim, K.; Park, Y.M.; Hong, C.S. Machine Learning based edge assisted UAV computation offloading for data analyzing. In Proceedings of the IEEE International Conference of Information Networking (ICOIN), Barcelona, Spain, 7–10 January 2020; pp. 117–120.
102. Wang, S.; Chen, M.; Yin, C.; Saad, W.; Hong, C.S.; Cui, S.; Poor, H.V. Federated learning for task and resource allocation in wireless high altitude balloon networks. *arXiv* **2020**, arXiv:2003.09375. [\[CrossRef\]](#)
103. Lim, W.Y.B.; Huang, J.; Xiong, Z.; Kang, J.; Niyato, D.; Hua, X.S.; Miao, C. Multi-Dimensional Contract-Matching for Federated Learning in UAV-Enabled Internet of Vehicles. In Proceedings of the GLOBECOM 2020—2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6. [\[CrossRef\]](#)
104. Ng, J.S.; Lim, W.Y.B.; Dai, H.N.; Xiong, Z.; Huang, J.; Niyato, D.; Miao, C. Joint Auction-Coalition Formation Framework for Communication-Efficient Federated Learning in UAV-Enabled Internet of Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 2326–2344. [\[CrossRef\]](#)
105. He, Y.; Zhai, D.; Huang, F.; Wang, D.; Tang, X.; Zhang, R. Joint Task Offloading, Resource Allocation, and Security Assurance for Mobile Edge Computing-Enabled UAV-Assisted VANETs. *Remote Sens.* **2021**, *13*, 1547. [\[CrossRef\]](#)
106. Zhang, Z.; Xie, X.; Xu, C.; Wu, R. Energy Harvesting-Based UAV-Assisted Vehicular Edge Computing: A Deep Reinforcement Learning Approach. In Proceedings of the 2022 IEEE/CIC International Conference on Communications in China (ICCC Workshops), Sanshui, Foshan, China, 11–13 August 2022; pp. 199–204. [\[CrossRef\]](#)
107. Hu, N.; Qin, X.; Ma, N.; Liu, Y.; Yao, Y.; Zhang, P. Energy-efficient Caching and Task offloading for Timely Status Updates in UAV-assisted VANETs. In Proceedings of the 2022 IEEE/CIC International Conference on Communications in China (ICCC), Sanshui, Foshan, China, 11–13 August 2022; pp. 1032–1037. [\[CrossRef\]](#)
108. Cheng, Y.; Xu, S.; Cao, Y.; He, Y.; Xiao, K. SBA-GT: A Secure Bandwidth Allocation Scheme with Game Theory for UAV-Assisted VANET Scenarios. In *Wireless Algorithms, Systems, and Applications (WASA 2022)*; Lecture Notes in Computer Science; Wang, L., Segal, M., Chen, J., Qiu, T., Eds.; Springer: Cham, Switzerland, 2022; Volume 13472. [\[CrossRef\]](#)
109. Zheng, K.; Sun, Y.; Lin, Z.; Tang, Y. UAV-assisted online video downloading in vehicular networks: A reinforcement learning approach. In Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC 2020-Spring), Antwerp, Belgium, 25–28 May 2020; pp. 1–5.
110. Samir, M.; Ebrahimi, D.; Assi, C.; Sharafeddine, S.; Ghayeb, A. Leveraging UAVs for Coverage in Cell-Free Vehicular Networks: A Deep Reinforcement Learning Approach. *IEEE Trans. Mob. Comput.* **2021**, *20*, 2835–2847. [\[CrossRef\]](#)

111. Wang, J.; Zhang, X.; He, X.; Sun, Y. Bandwidth Allocation and Trajectory Control in UAV-Assisted IoV Edge Computing Using Multiagent Reinforcement Learning. *IEEE Trans. Reliab.* **2023**, *72*, 599–608. [\[CrossRef\]](#)
112. Boussoufa-Lahlah, S.; Semchedine, F.; Bouallouche Medjkoune, L. Geographic routing protocols for Vehicular Ad hoc NETWORKS (VANETs): A survey. *Veh. Commun.* **2018**, *11*, 20–31. [\[CrossRef\]](#)
113. Abdel-Halim, I.T.; Fahmy, H.M.A. Prediction-based protocols for vehicular Ad Hoc Networks: Survey and taxonomy. *Computer. Netw.* **2018**, *130*, 34–50. [\[CrossRef\]](#)
114. Benamar, N.; Singh, K.D.; Benamar, M.; El Ouadghiri, D.; Bonnin, J.M. Routing protocols in vehicular delay tolerant networks: A comprehensive survey. *Comput. Commun.* **2014**, *48*, 141–158. [\[CrossRef\]](#)
115. Mangrulkar, R.; Atique, M. Routing protocol for delay tolerant network: A survey and comparison. In Proceedings of the 2010 International Conference on Communication Control and Computing Technologies, Nagercoil, Tamil Nadu, India, 7–9 October 2010; pp. 210–215.
116. Wu, C.; Yoshinaga, T.; Bayar, D.; Ji, Y. Learning for adaptive anycast in vehicular delay tolerant networks. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1379–1388. [\[CrossRef\]](#)
117. He, J.; Cai, L.; Pan, J.; Cheng, P. Delay analysis and routing for two-dimensional VANETs using carry-and-forward mechanism. *IEEE Trans. Mob. Comput.* **2017**, *16*, 1830–1841. [\[CrossRef\]](#)
118. Karthikeyan, L.; Deepalakshmi, V. Comparative study on non-delay tolerant routing protocols in vehicular networks. *Procedia Comput. Sci.* **2015**, *50*, 252–257.
119. Wheeb, A.H.; Nordin, R.; Samah, A.; Alsharif, M.H.; Khan, M.A. Topology-based routing protocols and mobility models for flying ad hoc networks: A contemporary review and future research directions. *Drones* **2021**, *6*, 9. [\[CrossRef\]](#)
120. Ajaz, F.; Naseem, M.; Ahamad, G.; Khan, Q.R.; Sharma, S.; Abbasi, E. Routing protocols for Internet of vehicles: A review. In *AI and Machine Learning Paradigms for Health Monitoring System*; Springer: Singapore, 2021; pp. 95–103.
121. Le, M.; Park, J.-S.; Gerla, M. UAV assisted disruption tolerant routing. In Proceedings of the MILCOM 2006—2006 IEEE Military Communications Conference, Washington, DC, USA, 23–25 October 2006; IEEE: New York, NY, USA, 2006; pp. 1–5.
122. Di Maio, A.; Palattella, M.; Engel, T. Performance Analysis of MANET Routing Protocols in Urban VANETs. *Ad Hoc Mob. Wirel. Netw.* **2019**, *11803*, 432–451.
123. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
124. Boyan, J.; Littman, M. Packet routing in dynamically changing networks: A reinforcement learning approach. *Adv. Neural Inf. Process. Syst.* **1993**, *6*, 671–678.
125. Khodayari, S.; Yazdanpanah, M.J. Network routing based on reinforcement learning in dynamically changing networks. In Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence ICTAI'05, Hongkong, China, 14–16 November 2005; p. 366.
126. Srinidhi, N.N.; Sagar, C.S.; Shreyas, J.; SM, D.K. An improved PROPHET-Random forest based optimized multi-copy routing for opportunistic IoT networks. *Internet Things* **2020**, *11*, 100203.
127. Nadarajan, J.; Kaliyaperumal, J. QOS aware and secured routing algorithm using machine intelligence in next generation VANET. *Int. J. Syst. Assur. Eng. Manag.* **2021**. [\[CrossRef\]](#)
128. Luo, L.; Sheng, L.; Yu, H.; Sun, G. Intersection-Based V2X Routing via Reinforcement Learning in Vehicular Ad Hoc Networks. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 5446–5459. [\[CrossRef\]](#)
129. An, C.; Wu, C.; Yoshinaga, T.; Chen, X.; Ji, Y. A Context-Aware Edge-Based VANET Communication Scheme for ITS. *Sensors* **2018**, *18*, 2022. [\[CrossRef\]](#) [\[PubMed\]](#)
130. Jafarzadeh, O.; Dehghan, M.; Sargolzaey, H.; Esnaashari, M.M. A Model-Based Reinforcement Learning Protocol for Routing in Vehicular Ad hoc Network. *Wirel. Pers. Commun.* **2022**, *123*, 975–1001. [\[CrossRef\]](#)
131. Wu, J.; Fang, M.; Li, H.; Li, X. RSU-Assisted Traffic-Aware Routing Based on Reinforcement Learning for Urban Vanets. *IEEE Access* **2020**, *8*, 5733–5748. [\[CrossRef\]](#)
132. Bi, X.; Gao, D.; Yang, M. A Reinforcement Learning-Based Routing Protocol for Clustered EV-VANET. In Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 12–14 June 2020; pp. 1769–1773. [\[CrossRef\]](#)
133. Zhang, D.; Zhang, T.; Liu, X. Novel self-adaptive routing service algorithm for application in VANET. *Appl. Intell.* **2019**, *49*, 1866–1879. [\[CrossRef\]](#)
134. Khan, M.U.; Hosseinzadeh, M.; Mosavi, A. An Intersection-Based Routing Scheme Using Q-Learning in Vehicular Ad Hoc Networks for Traffic Management in the Intelligent Transportation System. *Mathematics* **2022**, *10*, 3731. [\[CrossRef\]](#)
135. Rahmani, A.M.; Naqvi, R.A.; Yousefpoor, E.; Yousefpoor, M.S.; Ahmed, O.H.; Hosseinzadeh, M.; Siddique, K. A Q-Learning and Fuzzy Logic-Based Hierarchical Routing Scheme in the Intelligent Transportation System for Smart Cities. *Mathematics* **2022**, *10*, 4192. [\[CrossRef\]](#)
136. Fuertes, D.; del-Blanco, C.R.; Jaureguizar, F.; Navarro, J.J.; García, N. Solving routing problems for multiple cooperative Unmanned Aerial Vehicles using Transformer networks. *Eng. Appl. Artif. Intell.* **2023**, *122*, 106085. [\[CrossRef\]](#)
137. Wang, X.; Fu, L.; Cheng, N.; Sun, R.; Luan, T.; Quan, W.; Aldubaikhy, K. Joint Flying Relay Location and Routing Optimization for 6G UAV-IoT Networks: A Graph Neural Network-Based Approach. *Remote Sens.* **2022**, *14*, 4377. [\[CrossRef\]](#)
138. Hussain, S.; Sami, A.; Thasin, A.; Saad, R.M. AI-Enabled Ant-Routing Protocol to Secure Communication in Flying Networks. *Appl. Comput. Intell. Soft Comput.* **2022**, *2022*, 3330168. [\[CrossRef\]](#)

139. Wang, W.; Liu, Y.; Srikant, R.; Ying, L. 3M-RL: Multi-Resolution, Multi-Agent, Mean-Field Reinforcement Learning for Autonomous UAV Routing. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 8985–8996. [\[CrossRef\]](#)
140. Sliwa, B.; Schuler, C.; Patchou, M.; Wietfeld, C. PARRoT: Predictive ad hoc Routing fueled by reinforcement learning and trajectory knowledge. *arXiv* **2020**, arXiv:2012.05490.
141. He, C.; Liu, S.; Han, S. A Fuzzy Logic Reinforcement Learning-Based Routing Algorithm For Flying Ad Hoc Networks. In Proceedings of the 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA, 17–20 February 2020; pp. 987–991. [\[CrossRef\]](#)
142. Liu, J.; Wang, Q.; He, C.; Xu, Y. ARdeep: Adaptive and Reliable Routing Protocol for Mobile Robotic Networks with Deep Reinforcement Learning. In Proceedings of the 2020 IEEE 45th Conference on Local Computer Networks (LCN), Sydney, NSW, Australia, 16–19 November 2020; pp. 465–468. [\[CrossRef\]](#)
143. Yang, Q.; Jang, S.J.; Yoo, S.J. Q-Learning-Based Fuzzy Logic for Multi-objective Routing Algorithm in Flying Ad Hoc Networks. *Wirel. Pers Commun.* **2020**, *113*, 115–138. [\[CrossRef\]](#)
144. Roh, B.-S.; Han, M.-H.; Ham, J.-H.; Kim, K.-I. Q-LBR: “Q-Learning Based Load Balancing Routing for UAV-Assisted VANET”. *Sensors* **2020**, *20*, 5685. [\[CrossRef\]](#)
145. Jiang, S.; Huang, Z.; Ji, Y. Adaptive UAV-Assisted Geographic Routing With Q-Learning in VANET. *IEEE Commun. Lett.* **2021**, *25*, 1358–1362. [\[CrossRef\]](#)
146. He, Y.; Zhai, D.; Jiang, Y.; Zhang, R. Relay Selection for UAV-Assisted Urban Vehicular Ad Hoc Networks. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 1379–1383. [\[CrossRef\]](#)
147. Shah, S.A.A.; Illanko, K.; Fernando, X. Deep Learning Based Traffic Flow Prediction for Autonomous Vehicular Mobile Networks. In Proceedings of the 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), Norman, OK, USA, 27 September–28 October 2021; pp. 1–5. [\[CrossRef\]](#)
148. Ali Shah, S.A.; Fernando, X.; Kashef, R. Improved Vehicular Congestion Classification using Machine Learning for VANETs. In Proceedings of the 2024 IEEE International Systems Conference (SysCon), Montreal, QC, Canada, 15–18 April 2024; pp. 1–8. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.