IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 July 2023, accepted 31 July 2023, date of publication 4 August 2023, date of current version 10 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3302250

SURVEY

Machine Learning Empowered Emerging Wireless Networks in 6G: Recent Advancements, Challenges and Future Trends

HAFIZ MUHAMMAD FAHAD NOMAN[1], (Graduate Student Member, IEEE), EFFARIZA HANAFI[1], (Member, IEEE), KAMARUL ARIFFIN NOORDIN[1], (Senior Member, IEEE), KAHARUDIN DIMYATI[1], (Member, IEEE), MHD NOUR HINDIA[1], (Member, IEEE), ATEF ABDRABOU[2], (Member, IEEE), AND FAIZAN QAMAR[3], (Member, IEEE)

[1]Advanced Communication Research and Innovation (ACRI), Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia
[2]Department of Electrical and Communication Engineering, College of Engineering, United Arab Emirates University, Al Ain, United Arab Emirates
[3]Centre for Cyber Security, Faculty of Information Science and Technology (FTSM), Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor 43600, Malaysia

Corresponding authors: Kamarul Ariffin Noordin (kamarul@um.edu.my) and Atef Abdrabou (atef.abdrabou@uaeu.ac.ae)

This work was supported by United Arab Emirates University under Grant 12N107.

ABSTRACT The upcoming 6G networks are sixth-sense next-generation communication networks with an ever-increasing demand for enhanced end-to-end (E2E) connectivity towards a connected, sustainable world. Recent developments in artificial intelligence (AI) have enabled a wide range of novel technologies through the availability of advanced machine learning (ML) models, large datasets, and high computational power. In addition, intelligent resource management is a key feature of 6G networks that enables self-configuration and self-healing by leveraging the parallel computing and autonomous decision-making ability of ML techniques to enhance energy efficiency and computational capacity in 6G networks. Consequently, ML techniques will play a significant role in addressing resource management and mobility management challenges in 6G wireless networks. This article provides a comprehensive review of state-of-the-art ML algorithms applied in 6G wireless networks, categorized into learning types, including supervised and unsupervised machine learning, Deep Learning (DL), Reinforcement Learning (RL), Deep Reinforcement Learning (DRL) and Federated Learning (FL). In particular, we review the ML algorithms applied in the emerging networks paradigm, such as device-to-device (D2D) networks, vehicular networks (Vnet), and Fog-Radio Access Networks (F-RANs). We highlight the ML-based solutions to address technical challenges in terms of resource allocation, task offloading, and handover management. We also provide a detailed review of the ML techniques to improve energy efficiency and reduce latency in 6G wireless networks. To this end, we identify the open research issues and future trends concerning ML-based intelligent resource management applications in 6G networks.

INDEX TERMS 6G, D2D communication, energy efficiency, machine learning, resource management.

I. INTRODUCTION

Wireless communication systems have evolved to the recent fifth-generation (5G) networks, and upcoming 6G networks are anticipated to establish unlimited connectivity for a multi-layered and a large number of smart terminals [1]. Despite recent developments in 5G networks to support massive machine-type communication (mMTC), enhanced mobile broadband (eMBB), and ultra-reliable and low-latency communications (URLLC), it is foreseen that the 6G networks will be able to provide various emerging services, including end-to-end connected autonomous

The associate editor coordinating the review of this manuscript and approving it for publication was Eyuphan Bulut.

systems, zero-touch cognitive networks, Holographic-Type Communication (HTC), Tactile Internet, etc. [2], [3], [4]. These emerging applications with diverse requirements such as peak data rate of 1 Tbps, 99.99999% reliability, and latency ranging between 10-100 microseconds are beyond 5G communication specifications under International Mobile Telecommunications-2030 (IMT-2030) standards for 6G vision [5]. Henceforth, 6G shows its potential to offer higher bandwidth (THz band) and a promising data rate with lower latency and more reliability compared to 5G. It will not only improve the quality of service (QoS) and the quality of experience (QoE) for end-users but also offer a low-cost and sustainable communication infrastructure for 6G wireless networks to facilitate smart applications, including augmented and virtual reality, contactless payment, holographic projection, Internet of Senses (IoS), flying base stations, etc. [6], [7]. Furthermore, to meet the diverse requirements of 6G, artificial intelligence (AI) is envisioned to enable autonomous systems with distributed learning models [8]. These AI-enabled 6G networks are intended to automate network processes, analyze big data to make smart decisions and realize intelligent edge, fog, and cloud nodes with the ultimate goal of achieving seamless E2E connectivity that cannot be achieved with existing 5G standards [9].

Similarly, industry X.0 will also make 6G networks more dynamic, heterogeneous, and ultra-dense by unleashing new energy into industrial products. Consequently, heterogeneous devices in 6G networks must adhere to different QoS requirements for intelligent deployment of limited network resources (radio, computation, energy, etc.). For example, seamless connectivity and real-time data transfer between autonomous vehicles necessitate an ultra-reliable and low-latency network. The same physical infrastructure is anticipated to simultaneously support customers' demands for real-time video streaming and multiple entertainment applications utilizing extended reality (augmented/virtual reality), telemedicine, etc. [6], [7]. These incipient wireless network applications necessitate network services with a variety of performance characteristics such as resource allocation, task offloading, mobility management, energy efficiency, and latency minimization, posing fundamental technical challenges for intelligent resource management in 6G networks.

There is a tremendous need to address the aforementioned challenges for optimal utilization of network resources in the design guidelines of 6G networks. Furthermore, resource management is crucial for information exchange between vehicles, infrastructure, D2D connections, extended reality in teleoperation, low-latency transmission of safety-related and health alert messages, high-precision map navigations, etc. Particularly, it may be challenging to simultaneously meet reliability, computing efficiency, high data rate, low latency, and energy-efficient communication. Consequently, it is necessary to formulate joint optimization problems for resource management issues such as radio resource allocation, user association, power allocation, spectrum management, and computation offloading to meet various demands in 6G wireless applications.

In wireless networks, traditional resource management problems are addressed with suboptimal or heuristic optimization methods. On the other hand, the joint optimization problems for energy efficiency, resource allocation, and computation efficiency in upcoming 6G networks are NP-hard due to mixed integer nonlinear programming (MINLP). These NP-hard problems could not be solved using global optimization algorithms, e.g. the branch-and-bound algorithm, and sub-optimal or heuristic algorithms, due to their exponential computation complexity and difficulty quantifying and controlling the performance gaps to the optimal solution. To balance the computational complexity and performance gap of solving NP-hard problems, machine learning algorithms (a subset of artificial intelligence) have emerged as an alternative approach to heuristics and brute-force algorithms. This trend has also motivated researchers to investigate machine learning methods for joint optimization problems in 6G wireless networks. Moreover, 6G networks are increasingly stringent in terms of robustness, resource efficiency, and reliability. Consequently, a paradigm shift in traditional resource management methods is required for the joint optimization of network resources using ML-empowered intelligent resource management in 6G networks [10].

Additionally, the large bandwidth of the THz spectrum range offers high-data rates, which act as the driving force in optimizing 6G wireless networks using ML techniques. ML techniques implemented in 6G networks are envisioned to achieve efficient spectrum management, automatic network configurations, and cognitive provisioning of services [11]. Further, the combination of machine learning and 6G networks can intelligently optimize network resources, real-time learning, and complex decision-making for modern autonomous systems. Moreover, as the evolution towards 6G networks proceeds, it will be imperative to balance the expected explosive growth of data traffic, integrating sensor-based services with real-time analytics, and massive network densification with the demand for global sustainability and fairness [12]. Consequently, ML-enabled 6G networks will be a fundamental component for the functioning of virtually all parts of society, and industries, fulfilling the communication needs of humans as well as intelligent machines.

The scope of this survey paper is to provide a detailed overview of the challenges and potential advantages related to integrating ML techniques with advanced technologies in 6G networks. It covers the ML-assisted resource management challenges to enhance the performance of 6G wireless networks while optimization of wireless resources to support ML applications is not covered in this survey [8], [13]. This article explains how the existing works developed ML-based solutions to tackle intelligent resource management in 6G wireless networks, categorized into ML types, including supervised ML and unsupervised ML, Deep Learning (DL),

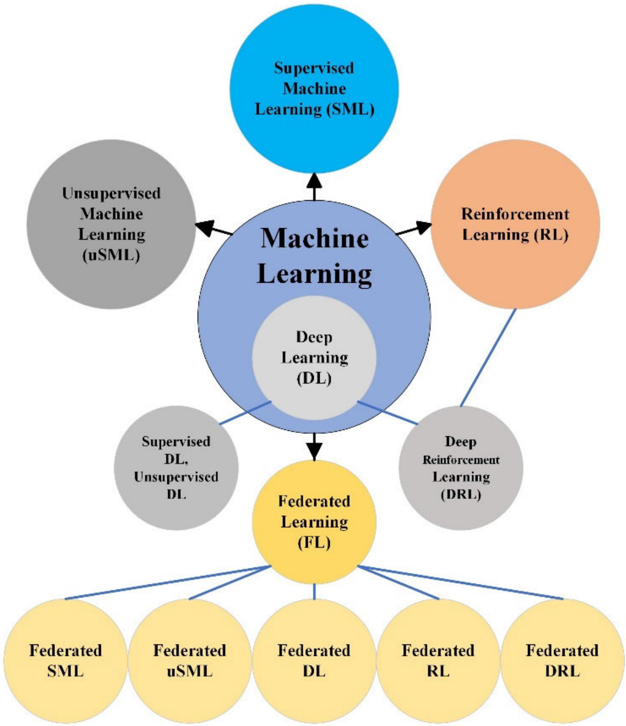


FIGURE 1. Taxonomy of ML techniques.

Reinforcement Learning (RL), Deep Reinforcement Learning (DRL) and Federated Learning (FL). Fig. 1 explains the taxonomy of ML types. These ML types have been classified into broader categories and subcategories presented in this article. Some types of ML are not mutually exclusive. For instance, deep learning (DL), a subset of ML, encompasses supervised DL and unsupervised DL as its subcategories. Conversely, supervised ML and unsupervised ML, as broader types of ML, do not encompass deep learning. Likewise, federated learning exhibits subcategories such as Federated SML, Federated uSML, Federated DL, Federated RL, and Federated DRL. Furthermore, DRL is a subcategory that is a combination of deep learning and reinforcement learning.

A state-of-the-art review of the existing literature is provided, focusing on the methodology, advantages, and limitations of each proposed ML algorithm concerning various network configurations, including D2D communication, vehicular networks, and Fog-Radio Access Networks. Furthermore, it categorizes the recent literature based on ML-empowered resource allocation, task offloading, energy efficiency, latency minimization, and handover management to provide guidelines for selecting the suitable category of an ML algorithm and highlight potential advantages, limitations and open issues concerning ML-assisted 6G wireless networks. Fig. 2 depicts the organization of the paper. In Table 1, the abbreviations used in this article are listed.

TABLE 1. List of abbreviations.

Acronym	Definition
5G Networks	Fifth-Generation Networks
6G Networks	Sixth-Generation Networks
AI	Artificial Intelligence
BDI	Belief Desire Intention
C-RAN	Cloud Radio Access Network
D2D	Device-to-Device
DL	Deep Learning
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
EE	Energy Efficiency
F-AP	Fog Access Point
FL	Federated Learning
F-RAN	Fog Radio Access Network
GRU	Gated Recurrent Unit
HO	Handover
IoT	Internet-of-Things
MDP	Markov Decision Process
ML	Machine Learning
QoE	Quality of Experience
QoS	Quality of Service
RA	Resource Allocation
RL	Reinforcement Learning
RNN	Recurrent Neural Network
SDN	Software-Defined Network
SML	Supervised Machine Learning
TO	Task Offloading
uSML	Unsupervised Machine Learning
Vnet	Vehicular Network
VR	Virtual Reality
V2X	Vehicle-to-Everything

II. RELATED WORK AND CONTRIBUTION

A. RELATED WORK

Intelligent resource management in heterogeneous and ultra-dense environments is the major challenge for improved energy efficiency, ultra-reliability, and overall low latency in 6G wireless networks. Recently, ML methods have emerged as a promising solution in 6G wireless networks to optimize network resources and improve system performance. Further, ML-enabled cross-layer network optimizations enhance the users' quality of experience. However, several optimization challenges of mixed integer nonlinear programming must be addressed to effectively implement ML methods in emerging networks such as D2D communication, vehicular networks, and Fog-RANs. Moreover, the selection of the optimal ML method for intelligent resource management problems such as resource allocation, task offloading, and handover management is also challenging. The researchers have started working on integrating ML with 6G networks to provide ubiquitous and reliable communication solutions with ML-enabled architectures. Nevertheless, many challenges and open issues still need further investigation.

Recently, many researchers reviewed ML implementation in wireless networks. For example, in [7], ML algorithms for effective IoT operations in 6G wireless networks are discussed. The authors summarized the different ML techniques to address resource allocation and energy efficiency in IoT scenarios. In [14], AI-enabled resource management beyond 5G networks is presented. This survey focuses on various physical layer design and optimization aspects, including channel measurements, modeling, estimation, and resource management. The authors in [15] reviewed the functions of autonomous and cooperative AI schemes for

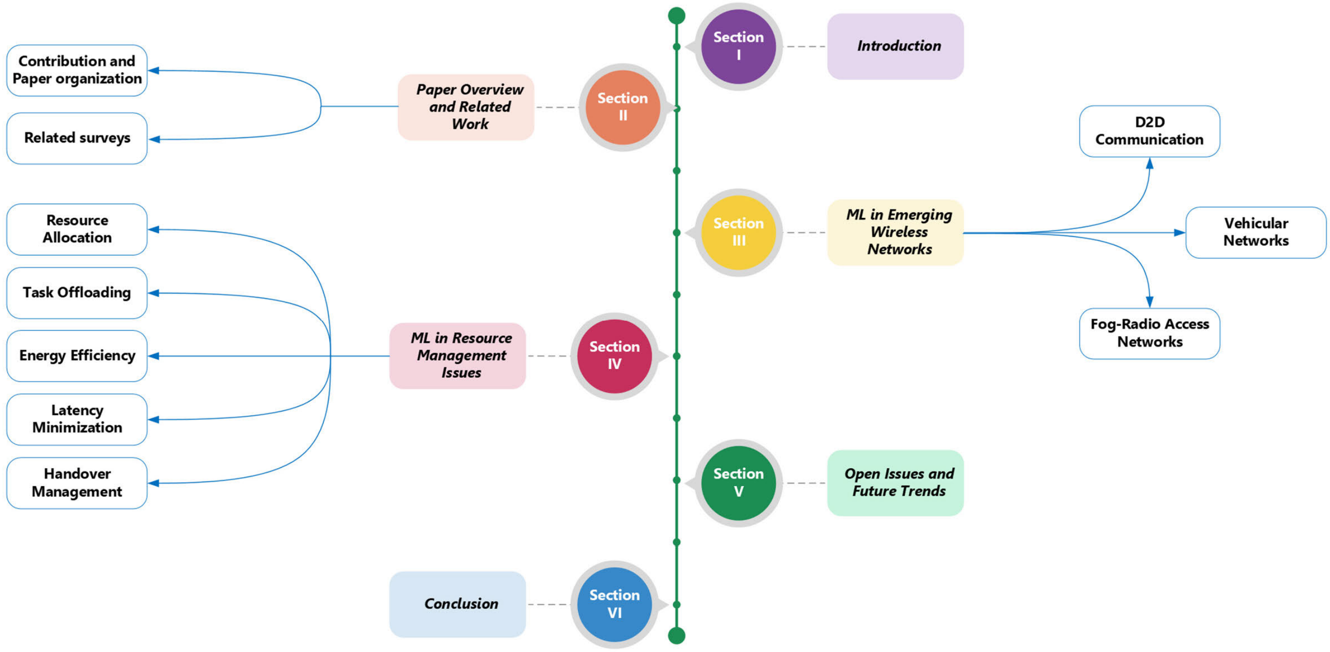


FIGURE 2. **Organization of the paper.**

resource allocation in 6G communication systems. In another work [16], the authors proposed a multi-layered design for 6G networks integrated with ML techniques. They discussed an integrated space–air–ground–underwater network (ISAGUN) as the potential core architecture of the 6G network. The references [7], [14], [15] and [16] proposed AI to enable intelligent wireless networks without considering how it can systematically sense data from environments, analyze the collected data, and then apply the discovered knowledge to optimize network performance for 6G networks.

Deep Learning (DL) implementation in wireless networks has been reviewed in recent works [17], [18], [19]. In [17], a DL-based multi-level architecture is proposed for data-driven 6G networks. It summarizes the deep learning-based network service provisioning at the device, edge, and cloud levels. This survey considers massive IoT cellular technology, haptic communication technology, and wireless channel modeling in 6G networks. The authors in [18] discussed DRL models, their applications, and challenges in autonomous IoT systems. The survey focuses on specific types of DRL models applied in an IoT environment with three layers: perception, network, and application layer. It highlights the pros and cons of existing works classified into different layers of IoT environment under the general DRL model. In another work [19], the authors reviewed the DL-based works for the 6G radio access networks and highlighted the main steps of DL model deployment in O-RAN. Moreover, this survey also covers the ML system operations concept in O-RAN. Nevertheless, these surveys [17], [18], [19] mainly focused on deep learning methods in IoT and O-RAN applications without reviewing the emerging network applications in 6G including D2D communication, vehicular networks, and F-RANs.

In [20], handover management issues in 6G networks using DRL and FL are presented. Recent studies have covered the use of ML in F-RAN systems, including data fusion and massive MTC applications [21], [22]. Furthermore, ML algorithms applied in D2D communication and ML-assisted load balancing in heterogeneous networks are discussed in [23] and [24], respectively. Nevertheless, these works concentrate on DL, DRL, and FL algorithms without considering the supervised, unsupervised, and model-free RL algorithms.

Addressing the network optimization problems in wireless communication is essential in 6G networks. ML-enabled cross-layer network optimization and resource management have been addressed in recent works [21], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36]. In [25], the authors discussed the 6G vision in AI-enabled edge systems. It summarizes the edge learning models in the context of communication-efficient edge training. Furthermore, it presents a comprehensive review of the resource allocation for edge AI systems. A detailed survey of ML techniques at the application and infrastructure level is given in [26], [27], and [28]. These works focus on resource allocation, power allocation, spectrum management, and wireless channel modeling using ML techniques. Additionally, there have been recent surveys on the implementation of ML methods for resource allocation and task offloading [29], [30], [31], [32], [33], [34], [35], resource management optimizations for ultra-massive access under URLLC constraints [36], whereas ML-assisted energy efficient communication in D2D [21], edge, and fog computing networks [37], [38], and large scale energy-harvesting networks are discussed in [36].

These aforementioned works have made a significant contribution to the research in ML-assisted next-generation wireless networks in terms of resource allocation, task offloading, handover management, latency minimization, etc., which enhances QoS and QoE in D2D communication, vehicular networks, and F-RANs. However, many of these issues are addressed individually in surveys [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36] with a focus on specific types of network scenarios. For instance, references [23], [28], [33], [37] consider only D2D communication networks. Similarly, vehicular networks are discussed in [39] and references [22], [32], and [40] only focus on F-RANs. Moreover, the D2D and Vnet are jointly considered in [20], [27], [34], and [35]. On the other hand, references [21] and [38] cover both D2D and F-RANs network scenarios in the context of ML-assisted wireless networks.

To the best of our knowledge, survey papers [7], [37] are the only ones that discuss all ML types (SML, uSML, DL, RL, DRL, FL) with a focus on resource management for D2D networks. However, while these two surveys consider resource management in D2D networks, they ignore vehicular networks and Fog-RANs. Moreover, they also do not cover ML-assisted handover management, energy efficiency, and latency minimization techniques for D2D networks. Furthermore, joint optimizations of resource management issues are not included in the aforementioned surveys. Nevertheless, there is no single survey article that combines the use of all types of ML techniques in D2D, Vnet, and F-RANs together to address the technical challenges in resource management issues such as resource allocation, task offloading, energy efficiency, latency minimization, and handover management. Table 2 summarizes the comparison of the recent surveys with our paper.

B. CONTRIBUTION OF SURVEY

To fill this gap and to stimulate further research in ML-empowered intelligent resource management in 6G networks, we present a novel and comprehensive survey comprising recent work on the state-of-the-art ML types (SML, uSML, DL, RL, DRL, FL) in three network categories including D2D communication, Vnet, and F-RANs. We also provide a detailed discussion on ML applications to address technical challenges in terms of resource allocation, task offloading, and handover management. A comprehensive summary of ML techniques to improve energy efficiency and reduce latency in 6G wireless networks is also provided. Further, this survey presents a new perspective and classification to recently published literature with a focus on ML-assisted 6G wireless networks, which leads us to identify the open issues and challenges in intelligent resource management and handover management. Moreover, we propose several promising future research directions in design guidelines of ML- empowered 6G wireless applications.

The main contributions in this article are outlined as follows:

- Present an overview of state-of-the-art ML techniques, such as supervised (SML) and unsupervised machine learning (uSML), Deep Learning (DL), Reinforcement Learning (RL), Deep Reinforcement Learning (DRL), and Federated Learning (FL) for resource management applications in 6G networks.
- Discuss the implementation of ML algorithms in emerging wireless networks such as device-to-device (D2D) networks, vehicular networks (Vnet), and Fog-Radio Access Networks (F-RANs).
- Summarize the ML-assisted solutions, including the learning types, advantages, and limitations for resource allocation, task offloading, and handover management.
- Provide a comprehensive review of ML-assisted methods to improve network performance in terms of energy efficiency and latency minimization.
- Highlight the open issues, research challenges, and possible solutions with future research trends in the context of ML applications in design guidelines of 6G wireless applications.

C. PAPER ORGANIZATION

A conceptual diagram of the topics covered in this article is shown in Fig. 2. It depicts the section-wise organization of the paper. Section II highlights the related work with a summary of the recent survey papers that employ ML in wireless communication systems and an overview of the article. Section III discusses ML methods applied in emerging wireless networks including, D2D communication, vehicular networks, and F-RANs. Section IV summarizes the ML techniques for addressing technical challenges in resource allocation, task offloading, and handover management. It also provides a comprehensive survey of ML techniques for energy efficiency and latency minimization in wireless networks. In Section V, open issues and future research trends are highlighted. Finally, Section VI presents the conclusion of the paper.

III. ML APPLICATIONS IN EMERGING WIRELESS NETWORKS PARADIGM

ML-empowered 6G networks are intended to provide cross-layer network optimizations in emerging wireless networks paradigm, including D2D communication, Vnets, and F-RANs by analyzing the big data to make smart decisions and enable intelligent edge, fog, and cloud nodes to achieve seamless E2E connectivity. A multi-layered ML-enabled network architecture for 6G applications is proposed in Fig. 3. It envisions context-aware smart resource management, self-healing network configuration, computation offloading, and intelligent service provisioning using ML techniques.

The architecture can be broadly classified into four layers, including the sensing layer, data analytics layer, control layer, and application layer. The sensing layer considers

TABLE 2. A comparison of recent surveys exploiting ML techniques in wireless networks.

Ref	Machine Learning Type						Network Scenario			Resource Management Technical Challenges					Distributed Computing	
	SML	uSML	DL	RL	DRL	FL	D2D	Vnet	F-RAN	RA	TO	EE	Latency	HO	Edge	Fog
[7]	✓	✓	✓	✓	✓	✓	✓			✓						
[14]			✓						✓	✓			✓		✓	✓
[15]	✓	✓		✓	✓		✓			✓			✓			
[16]	✓	✓	✓	✓	✓								✓	✓	✓	
[17]	✓	✓	✓	✓	✓					✓		✓	✓		✓	
[18]					✓			✓		✓	✓	✓	✓		✓	✓
[19]			✓		✓	✓				✓				✓	✓	
[20]	✓	✓		✓		✓	✓	✓					✓	✓		
[21]			✓	✓			✓		✓	✓	✓	✓	✓		✓	✓
[22]	✓	✓	✓						✓	✓	✓		✓		✓	✓
[23]							✓			✓			✓		✓	
[24]	✓			✓	✓									✓		
[25]			✓	✓	✓	✓				✓		✓	✓		✓	
[26]	✓	✓	✓	✓		✓				✓		✓	✓			
[27]			✓	✓		✓	✓	✓		✓	✓				✓	
[28]	✓	✓	✓	✓			✓							✓	✓	
[29]			✓		✓	✓				✓			✓		✓	
[30]	✓	✓	✓	✓		✓				✓			✓		✓	
[31]	✓	✓	✓	✓						✓			✓			
[32]			✓	✓	✓				✓	✓	✓	✓				✓
[33]	✓	✓	✓	✓	✓		✓			✓		✓				
[34]				✓	✓		✓	✓		✓						
[35]		✓					✓	✓		✓			✓			
[36]	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓	✓	
[37]	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓			✓	
[38]			✓	✓	✓	✓	✓		✓	✓		✓	✓			✓
[39]	✓	✓	✓	✓		✓		✓					✓			
[40]					✓	✓			✓	✓			✓			✓
[41]	✓	✓	✓	✓											✓	✓
[42]	✓	✓	✓	✓						✓		✓	✓	✓		
This work	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

the sensing nodes in the form of different types of UEs or devices (cellular users, ground vehicles, UAVs, etc.) The data analytics layer comprises an intelligent cloud with smart data centres, computation servers, and storage. ML-enabled SDN controllers are the backbone of the control layer that is responsible for various functionalities, including perception, action, feedback, global model aggregation, etc. Intelligent application servers communicate with edge devices to aggregate local ML models with local data perception to deliver smart applications using D2D, V2X, and Fog-RAN communication scenarios.

Furthermore, the proposed architecture includes macro base stations (MBSs), micro base stations, (μBSs), and small base stations (SBSs) with multiple computation servers at different layers considering the heterogeneous and ultra-dense nature of 6G networks. It includes cloud servers, fog servers, edge servers (access points), and UEs. The MBSs are capable of communication across sub-6 GHz, mmWave, and THz frequency bands, while the μBSs and APs support small-cell connectivity utilizing mmWave and THz frequencies. The interconnections of MBSs, μBSs, and APs are facilitated through the deployment of fiber optic links in the backhaul

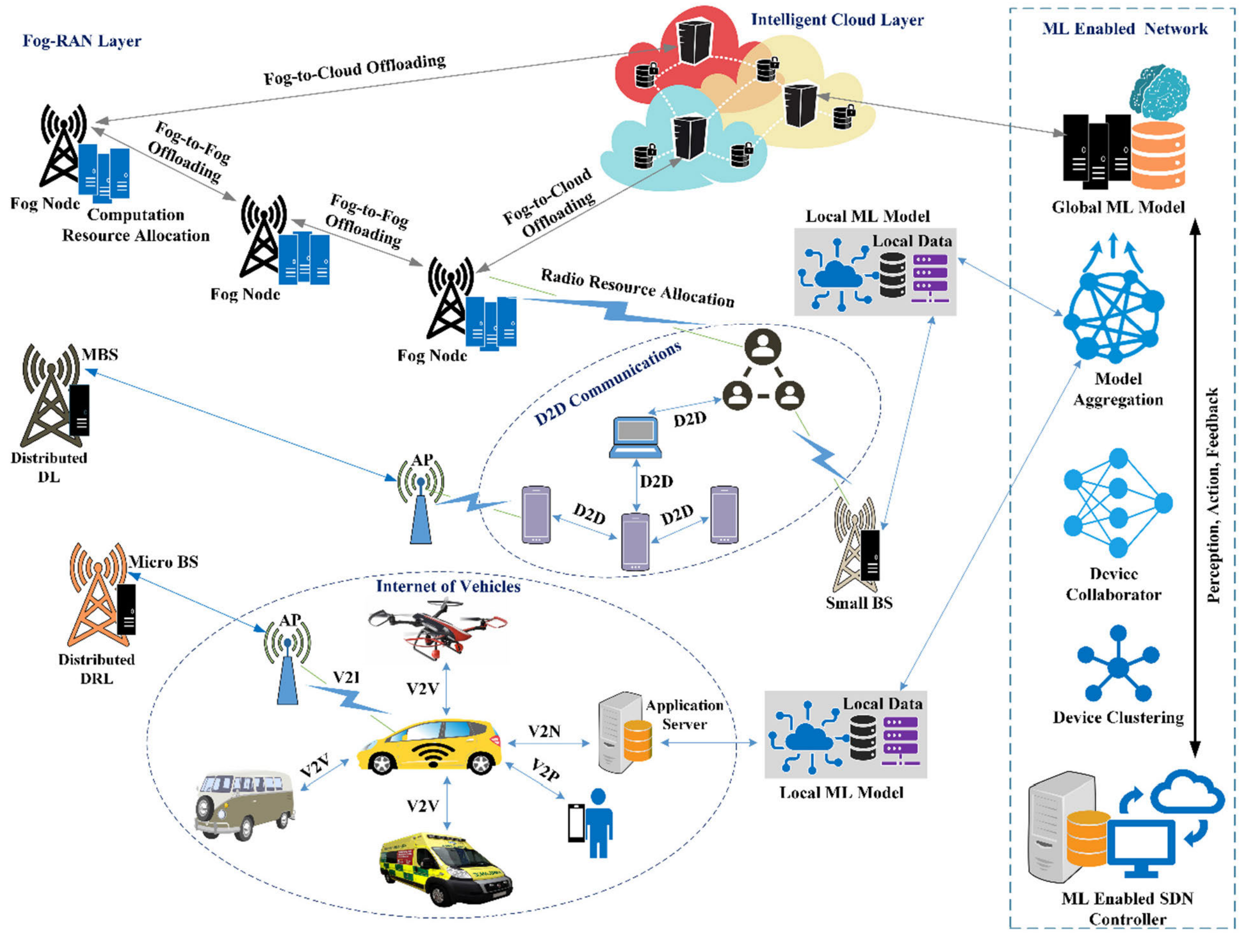


FIGURE 3. ML empowered emerging wireless networks architecture.

network infrastructure, enabling efficient, reliable, and low-latency transmission. These fiber optic links also play a vital role in facilitating interlayer communication within the network architecture. The transmitted data among MBSs, μBSs, and APs primarily consists of control signaling, coordination information, and routing protocols, which are crucial for orchestrating network operations, resource allocation, seamless handovers, and coverage coordination. Moreover, the optical network serves as a critical component for supporting the offloading and distribution of ML tasks across the APs and BSs, spanning multiple network layers.

The cloud and fog servers can participate in collaborative learning for task offloading. In this approach, an ML model can be trained and tested on the cloud or fog server using collaborative learning. Furthermore, the cloud servers can collaborate with multiple cloud servers on the same cloud layer or multiple fog servers at the fog layer. Henceforth, the cloud servers can communicate with the base stations in the fog layer to distribute the ML models among the fog servers to execute the task offloading using collaborative learning. In addition, fog servers can further distribute machine learning models to edge servers received from the cloud servers. Nevertheless, a latency-sensitive ML task may not be distributed among multiple fog servers by the cloud servers. Consequently, cloud servers execute such tasks to meet the QoS requirements and minimize the impact of communication latency and signaling overhead due to multi-layered architecture.

Furthermore, to train and test a particular ML model, the cloud, fog, and edge servers can use supervised, unsupervised, and reinforcement learning algorithms. Moreover, these servers can employ distributed deep learning and deep reinforcement learning for application-specific ML training. The proposed architecture is designed around these aspects to overcome the technical challenges associated with resource management and handover problems that will inevitably burden 6G networks. This section covers the implementation of ML algorithms in emerging wireless networks such as device-to-device (D2D) networks, vehicular networks (Vnet), and Fog-Radio Access Networks (F-RANs).

A. DEVICE-TO-DEVICE COMMUNICATION (D2D)

With the recent developments in machine learning methods, more research is turning to implement ML algorithms for addressing D2D communication challenges in 6G wireless networks. ML methods have the potential to overcome resource optimization problems in D2D connectivity. In 6G wireless networks, ML-assisted D2D communication offers a higher transmission rate, enhanced energy efficiency, and higher spectral efficiency. It reduces the mobile traffic load and improves the system's capacity in many applications, including public safety, offloading network traffic, multimedia services, and ubiquitous vehicular communications [43]. However, the joint optimization of resources (energy, computation, spectrum, etc.) and interference management in D2D communications require further investigations to overcome these challenges. Recent studies for ML-based D2D communication are reviewed in this section. We have categorized these recent works according to the type of ML (uSML, SML, DL, RL, DRL, and FL) and the application scenario in D2D communication, as shown in Fig. 4.

Supervised Machine Learning (SML): A mixed-mode clustering approach in D2D communication is proposed in [44]. It identifies whether the users communicate with cluster heads (CHs) using D2D connectivity or whether the users establish a direct communication link with eNB. ML models, including Support Vector Machine (SVM), Random Forest (RF), and Deep Neural Networks (DNN), are trained as binary classifiers. These ML models aim to maximize the throughput of a specific user in binary classification problems. The accuracy of the classification techniques and the performance of individual users is compared by measuring throughput, energy consumption, and fairness of throughput. However, the proposed approach considers only multimedia content sharing in a non-heterogeneous environment.

Unsupervised Machine Learning (uSML): In [45], D2D communication is discussed as an optimization problem exploiting distributed artificial intelligence techniques. The authors proposed Belief Desire Intention (BDI) intelligent agents with extended capabilities (BDIx) to consider every D2D node separately in an autonomous way without using the base station. The Distributed Artificial Intelligent System (DAIS) algorithm is used for transmission mode selection. The agent's role is determined by the beliefs and events it perceives. The weighted data rate (WDR) is proposed as the decision metric for transmission mode selection for each UE. Through relays and clusters, the WDR aids in calculating the optimum route to the BS. The proposed approach considers high spectral efficiency, better data rate, and low computational load. However, it is assumed that a BDIx agent always tends to select the unused resource blocks that thrust aside resource management and interference management.

ML-based resource management framework in D2D underlaying cellular network is investigated in [46]. It addresses the reliability requirement for D2D users and the high throughput requirement of cellular users under uncertain channel state information (CSI). A robust optimization strategy based on support vector clustering (SVC) is proposed to create the CSI uncertainty as a compact convex set. It estimates the uncertainty set from CSI samples using the group of samples with imperfect CSI. The CSI samples with correlation and asymmetries are grouped in a high-dimensional feature space. The authors proposed a bisection search-based algorithm that uses convex constraints for power allocation. The proposed method enhances the system throughput and convergence. However, inter-cell and intra-cell interference management are not addressed in throughput maximization.

Deep Learning (DL): DL techniques automatically extract the features from data with complex structures and inner correlations, making them a suitable candidate for predictions, classifications, or complex decisions in D2D communication [47], [48], [49], [50]. In [47], the authors proposed a framework for content caching in D2D communication using deep learning models with a combination of RNN and DNN. It maximizes the average D2D cache-hit ratio. However, the computation complexity of the proposed approach is not discussed. The authors of [48] considered DL models, including FFN, CNN, GRU, and LSTM, to forecast changes in received signal strength in D2D pairs. The proposed approach gives the minimum input-sample length required by the training model to achieve an optimum prediction performance. However, LSTM and GRU are not compared in terms of performance to determine which model is more appropriate under this application scenario.

In [49], the authors addressed the energy harvesting and spectrum management issues in D2D communication using DNN. The authors evaluated the transmitted power and power splitting ratio using a DNN-based algorithm and iteration-based approaches, including exhaustive search (ES) and gradient search (GS). The proposed method provides global optimality requiring less time. Nonetheless, the hybrid time-switching and power-splitting with the effect of D2D pairs require further investigation. In another study [50], DNN is used for CSI encoding and resource allocation in D2D cellular networks. This approach maintains the QoS of cellular users while maximizing the overall spectral efficiency. However, further performance evaluation of the proposed mechanism is required based on real-time experimental data.

Reinforcement Learning (RL): A mode selection mechanism in a D2D heterogeneous network is proposed in [51]. The process is divided into two phases: the first phase consists of D2D rapid clustering and identification of density peaks. In the second phase, the transmission mode selection is implemented using RL. It maximizes the overall throughput by exploiting the multi-agent learning (MARL) with Nash-Q-learning and the Wolf-PHC policies. The performance of both RL strategies is evaluated in comparison with the stochastic algorithm, Deep Q-learning, and the simple greedy algorithm. While Nash Q-learning converges more quickly in small-scale scenarios, the WoLF-PHC algorithm is less complex and needs less process-

ing space in large-scale scenarios. In [52], MARL based framework is presented for content delivery in the F-RAN system with D2D users. The authors proposed a decentralized cross-layer network coding-coalition formation (CLNC-CF) switch method that provides a stable F-AP and CE-D2D

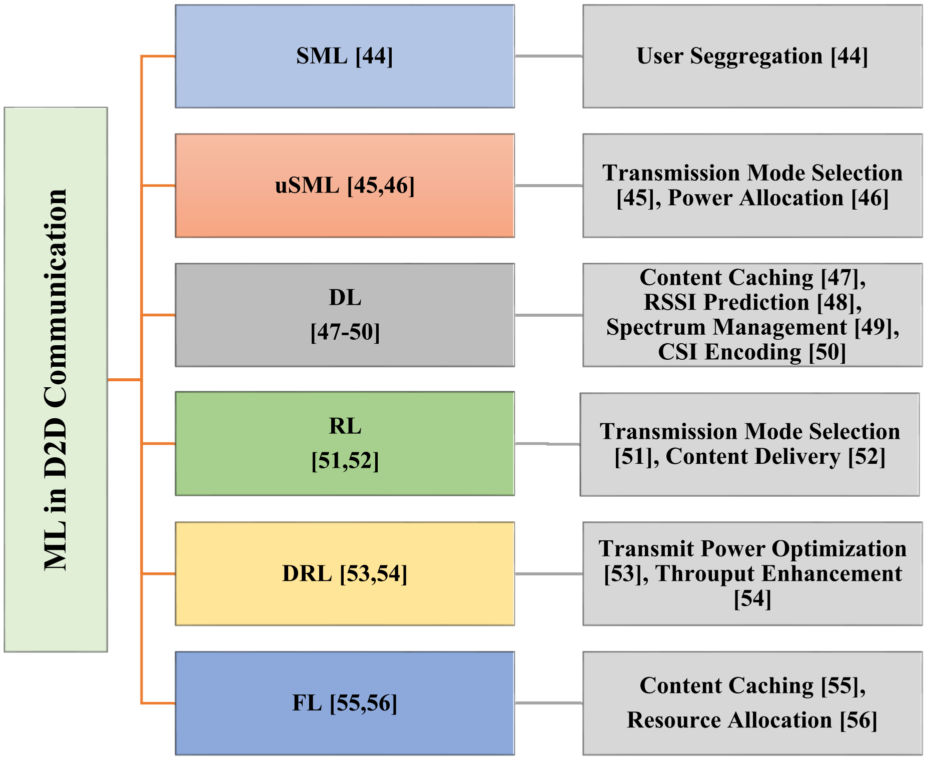


FIGURE 4. Recent work classification for ML techniques in D2D Communication.

user coalition. CLNC-CF and MARL solutions, collectively known as CLNC-CF-RL, enhance the sum rate, optimize the power level, and improve the transmission rates. However, the mobility of devices that may affect the system's power consumption is not considered.

Deep Reinforcement Learning (DRL): In [53], a DRL method is presented to minimize the transmit power for D2D users. The DRL algorithm is trained using periodic complex decision-making to minimize the loss function and achieve an optimum solution. The proposed approach has less computational complexity than ES. In [54], D2D-assisted deep Q-Learning is proposed for random access in mMTC networks. It considers a device clustering approach for devices and non-orthogonal multiple access (NOMA) mechanism to maximize the system throughput. A pre-clustering strategy is used to utilize the D2D links allowing the devices to run with a smaller Q-Table and, as a result, it improves convergence. Nonetheless, the proposed framework could be extended to include devices with variable target outage likelihood and spectral management requirements.

Federated Learning (FL): An intelligent and reliable D2D caching scheme is proposed in [55]. The authors presented a blockchain-empowered Deep Reinforcement Federated Learning (BDRFL) framework to ensure privacy and enhance data security in cache-enabled D2D networks. The proposed caching scheme shows significant performance to minimize the average latency with better convergence speed. However, congestion-aware data transfer and load balancing are not considered in cooperative D2D communication. In [56], the authors investigated resource allocation for D2D-assisted digital twin-edge networks. A Federated Reinforcement Learning-based approach is proposed to develop a decentralized global resource allocation strategy. It aims at optimizing both communication and power resources, ensuring high transmission rates for D2D links while maintaining efficient communication for cellular links. It also enables the digital twin edge network to achieve seamless connections using digital twins of user devices (UD-DT) and digital twins of access points (AP-DT). However, the cross-cell interference which may affect the system throughput is not considered.

Summary: Some of the recent works on ML implementation in D2D communication are outlined. In addition, the recent works are categorized according to the type of ML (SML, uSML, DL, RL, DRL, FL) and the application scenario in D2D communication, such as user segregation [44], transmission mode selection [45], [51], power allocation [46], content caching, and delivery [47], [52], RSSI prediction [48], spectrum management [49] CSI encoding [50], transmit power optimization [53], throughput enhancement [54], content caching [55], and resource allocation [56] as shown in Fig. 4. The proposed ML approach, an overview of the advantages and limitations of the recent works are provided in Table 3. The reviewed articles intended to enhance the throughput, spectral efficiency and reduce the computation complexity. However, there are several limitations to the current research, including load balancing, throughput fairness, and interference management, which need to be addressed and considered further.

B. VEHICULAR NETWORKS (VNET)

Vehicular networks are envisioned as fundamental communication platforms for various intelligent transportation systems (ITS), including traffic management and surveillance, incident management, video-based traffic condition detection, autonomous transit control, electronic toll collection, and road safety warnings [57]. The emerging ITS applications enhance the traffic system's sustainability, efficiency, and safety. However, these innovative applications require intensive computations and real-time data transfer [58]. Vehicle-to-Everything (V2X) networks, or autonomous vehicle technologies, have evolved into Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), Vehicle-to-Pedestrian (V2P), and Vehicle-to-Sensors (V2S) connectivity [59]. The V2X networks aim to provide not only infotainment and multimedia-based applications, such as augmented reality, autonomous driving, and crowd sensing but also improve road safety-related critical applications. It requires reliable and latency-aware communications through ubiquitous connectivity to V2X servers using V2I links and real-time safety-critical message transmission to nearby vehicles through V2V or V2S communications. This section comprises recent studies that employ ML techniques in vehicular networks. The presented work is classified based on the type of ML (uSML, SML, DL, RL, DRL, and FL) and the application scenario in vehicular networks, as shown in Fig. 5.

Supervised Machine Learning (SML): In [60], supervised ML techniques, including Naive Bayes, Support Vector Machine (SVM), K Nearest Neighbor (KNN), and Random Forest, are developed to forecast congestion warnings in heterogeneous vehicular communication. This method aims at enhancing the stability of the network and the classification of dynamic congestion. The suggested method minimizes the packet loss ratio and average delay. It also improves the system throughput. However, the robustness of the proposed model could be enhanced by considering more input features such as mobility, the complexity of scenarios, and the number of RBs.

Unsupervised Machine Learning (uSML): The authors in [61] addressed mobility management in fog-assisted vehicular networks. AI-enabled multi-layered F-RAN architecture is developed for intercity communication. The first level has a cloud layer offering access to the lower layers. The service layer is responsible for considering speed (fading, interference, and shadowing), reliability (packet-loss ratio and latency), and mobility (throughput and RSSI). In contrast, the vehicular layer is responsible for integrating the vehicles using V2V, V2I, and vehicle-to-roadside units (RSUs) communication links. The proposed framework provides an information exchange mechanism between various heterogeneous technologies to develop a hybrid F-RAN-based vehicular environment. However, the proposed approach does not consider cooperative communication between vehicular and fog servers.

In [62], a task offloading solution for vehicular networks under a hybrid fog and cloud model is proposed. It considers the heterogeneous characteristics of vehicular fog, fog servers, and the central cloud. The authors proposed the probabilistic task offloading (PTO) algorithm using the alternating direction method of multipliers (ADMM) and particle swarm optimization (PSO). It addresses the PTO problem as a decomposition coordination process. An iterative process based on three iterations is proposed to update the ADMM algorithm. The proposed approach reduces energy consumption and improves the average execution delay. Nevertheless, the proposed model does not take into account the packet loss and interference that may impact the lower communication issues.

Deep Learning (DL): In [63], a DL-based method is proposed for resource allocation in cyber twin-driven connected vehicles. It customizes the complexity of the DL model using efficient feature selection under the availability constraints of the computing resources. The proposed framework is validated through simulations for real-time estimation of battery status in electric vehicles. A DNN is proposed to monitor the real-time battery status by learning of relations between battery features and the system on chip (SoC). The customized model approach has the lowest computational overhead as the feature selection phase is executed in one round. The proposed method leverages adaptive AI modeling to optimize the model accuracy and computational complexity in cyber twin-driven connected vehicles. However, cooperative learning that may reduce the overall computational overhead is not addressed.

In [64], DL models are discussed for resource allocation in SDN-enabled vehicular networks. These models include CNN, DNN, and LSTM-based architecture. The proposed approach optimizes allocation time and accuracy while allocating the resources to different network slices. Nonetheless, the proposed approach could be extended to achieve better throughput using Gated Recurring Units and bidirectional LSTM.

Reinforcement Learning (RL): In [65], the authors used reinforcement learning to investigate resource management in IoV. In highly dynamic V2X scenarios, the proposed method ensures reliability with lower latency using software-defined vehicular architecture (SDV-F). It optimizes the resource utilization and distribution of the traffic and reduces the average delay. However, considering the distributed load-balancing strategy could extend the proposed model to improve networks' survivability and reliability.

Deep Reinforcement Learning (RL): In [66], the authors proposed DRL for mode selection and content caching in F-RAN slicing. The RAN slice instances are exploited to serve the UEs for the hotspot configuration and V2I application. The authors proposed a traditional Epsilon-Greedy

TABLE 3. Recent work summary of ML techniques in D2D communication.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/ Future work
[44]	SML	SVM, RF, ANN	A user segregation scheme targeting D2D clustering in a cellular network.	Enhance throughput and fairness of the throughput, decrease energy consumption.	The impact of packet loss and interference in D2D pairs is not addressed.
[45]	uSML	BDI Intelligent agents	Transmission mode selection in D2D communication.	Enhance spectral efficiency by 30%, reduce power consumption, and computational load.	The unused resource block selection by the BDIx agent ignores the impact of interference.
[46]		SVC	D2D power allocation under uncertain CSI.	System model robustness, enhance system throughput.	The inter-cell and intra-cell interference management are not addressed in throughput maximization.
[47]	DL	RNN, DNN	Collaborative filtering-based content caching in D2D communication.	Maximize the average D2D cache-hit ratio.	The computation complexity of the proposed method is not discussed.
[48]		FFN, CNN, GRU, LSTM	Predicting the received signal strength variations in D2D networks.	Reduce the input-sample length to enhance target prediction accuracy.	Prediction accuracy analysis of LSTM and GRU is not considered.
[49]		DNN	Energy harvesting and spectrum management in D2D network.	Enhance spectral efficiency, reduce time complexity.	The hybrid time-switching and power-splitting with the effect of D2D pairs are not addressed.
[50]		DNN	Resource allocation and CSI encoding in D2D communication.	Enhance spectral efficiency, improve cellular users' QoS.	The proposed scheme does not address the power consumption and computational load.
[51]	RL	WoLF-PHC, Nash Q-Learning	Hybrid transmission mode selection in D2D HetNets for VR broadcasting.	Enhance the system throughput. Reduce resource cost.	The spectral efficiency is not addressed in throughput maximization.
[52]		Multi-agent RL	Content delivery in the F-RAN system with D2D users.	Enhance the sum rate, optimize the power level, and improve the transmission rate.	The mobility of devices that may affect the system's power consumption is not considered.
[53]	DRL	Q-Learning	Transmit power optimization for cellular and D2D users.	Less computational complexity as compared to ES used for the global optimal solution.	Constraints on system fairness like proportionate fairness are not addressed.
[54]		Q-Learning	Throughput improvement in massive MTC networks with D2D clustering.	Improve the convergence speed, reduce the Q-Table size and computational complexity.	The spectral efficiency and devices with different target outage probability requirements are not considered.
[55]	FL	F-DRL	Reliable and secure cache-enabled D2D communication.	Minimize the average latency, improve convergence speed, enhance data privacy and security.	Load balancing and congestion-controlled data transfer could be considered in cooperative D2D communication.
[56]		F-RL	Resource allocation in D2D-enabled digital twin edge network.	Maximize the sum capacity of D2D users, reduce the outage probability.	The cross-cell interference which may affect the system throughput is not considered.

algorithm to balance the exploration and exploitation rate. A DQN model is used to learn the content's popularity using the current state, selected action, associated reward, and subsequent state. The target value function and approximate value function are trained using the experience samples. The proposed solution increases learning efficiency, particularly considering the large state and action spaces. It also provides a framework to obtain network optimizations in autonomous decision-making by exchanging the minimum information. However, the issues like robustness to model drift, handling outliers, and safe learning are not considered.

In [67], a DQN-based framework for radio resource allocation in vehicular networks is proposed. It optimizes the transmission success rate using the generalized closed-loop. In this approach, an anchor node exploits the status of radio resource utilization using the recent uplink of the vehicle. Nonetheless, data transmission success rates could be improved by allocating both radio and network resources to each downlink task. In [68], DRL is proposed to address resource management and multi-vehicle task offloading in fog-assisted vehicular networks. It exploits the contract theory to examine an incentive mechanism where vehicles contribute computation resources and receive rewards from IoV in exchange. It reduces the computation complexity to take offloading decisions faster through DQN and improves the QoS. However, a greater number of vehicles causes the DQN output value space to increase, which reduces the performance of the system.

In [69], the authors proposed a DRL method for blockchain-enabled content caching in vehicular networks.

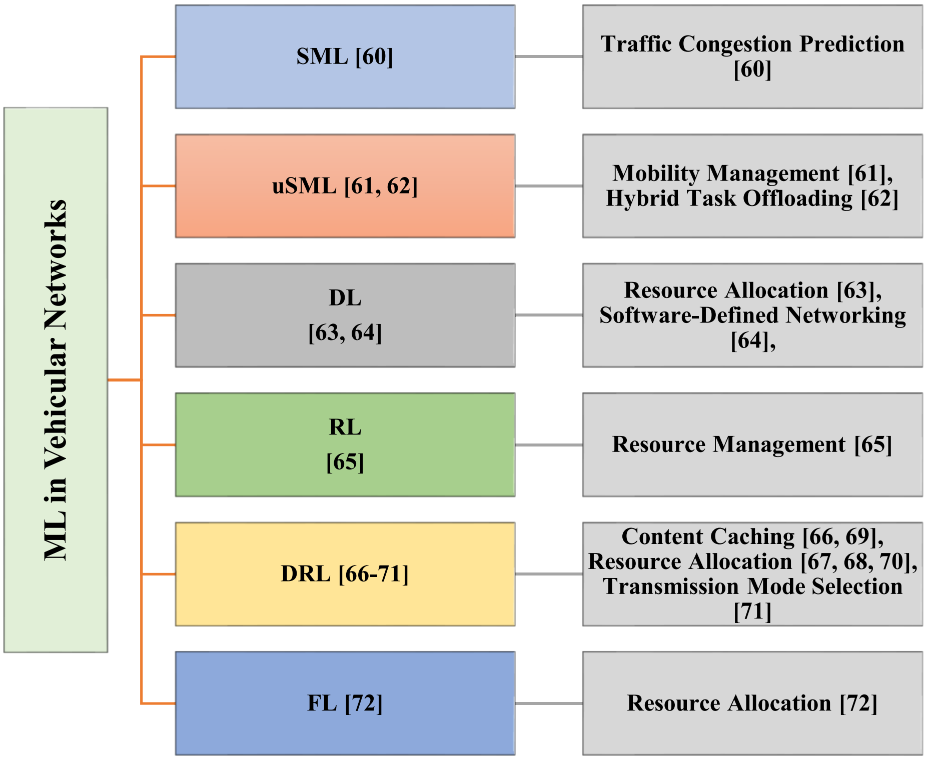


FIGURE 5. Recent work classification for ML techniques in vehicular networks.

The deep deterministic policy gradient (DDPG) is exploited to address the V2V content caching problem while using blockchain to maintain the security and privacy of content caching. However, content transmission latency is not considered in the proposed V2V communication.

In [70], DRL-based autonomous resource provisioning in virtualized D2D-based vehicular networks is discussed. In the proposed scheme, the resources of the mobile virtual network operators are adjusted by DRL agents to ensure an efficient resource allocation for interslice configuration and resource aggregation for D2D pairs. The proposed method balances the resource allocation and the QoS satisfaction level in addition to enhancing the overall system throughput. Nevertheless, the proposed framework is only validated for D2D-assisted virtualized V2V communication.

The authors in [71] proposed a DRL method for transmission mode selection and resource allocation in the V2X network. It addresses QoS heterogeneity issues and reliability requirements in V2V mode. It optimizes transmission mode selection, RB allocation, and power allocation to enhance the combined capacity of V2V links considering latency and reliability constraints. In the proposed approach, graph-based vehicle clustering is used on a large timescale, while federated learning is used on a small timescale for training the DRL models. However, the impact of packet priority on QoS enhancement and communication link quality is not considered.

Federated Learning (FL): In [72], resource allocation for V2V communication using Federated Multi-agent Deep Reinforcement Learning is discussed. A joint optimization problem of channel selection and transmit power control is formulated using Double Deep Q Network (D3QN). It exploits the V2V agent for decentralized channel selection and power allocation while maximizing the sum rate of cellular users and V2V packet delivery rate. However, the adaptive switching between centralized and decentralized channel access for the V2V pair is not considered.

Summary: Some of the recent works using ML techniques in vehicular networks (Vnets) are presented. The proposed ML approach and an overview of the advantages and limitations of the recent works are provided in Table 4. In addition, the recent works are categorized according to the type of ML and the application scenario in Vnet, such as traffic congestion prediction [60], mobility management [61], hybrid task offloading [62], resource management [63], [65], [67], [68], [70], software-defined networking [64], content caching [66], [69], transmission mode selection [71], resource allocation [72] as illustrated in Fig. 5. These recent works focused on improving the QoS of vehicular application, enhancing learning accuracy, and reduce computation complexity and the average end-to-end delay. Nonetheless, there are several limitations to the current research, including co-tier interference, cooperative communication between vehicular servers, and the impact of packet loss which need more attention and consideration.

TABLE 4. Recent work summary of ML techniques in vehicular networks.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/ Future work
[60]	SML	Naive Bayes, SVM, KNN, RF	Network congestion prediction for heterogeneous vehicular networks.	Improve throughput by lowering the average delay and packet loss ratio.	The mobility, complexity of scenarios, and number of RBs are not considered in the proposed model.
[61]	uSML	RIMMA	Mobility management for intra-vehicular networks.	Optimize the throughput, channel gain, and scalability during vehicular communication.	Cooperative communication between the vehicular server and the fog server is not addressed.
[62]		PSO	Offloading in vehicular networks using a hybrid fog/cloud service.	Reduce energy consumption, improve the average execution delay.	The impact of interference and packet loss is not considered in the proposed model.
[63]	DL	DNN	Resource allocation in cyber twin-driven connected vehicles.	Optimize model accuracy and computational complexity.	The cooperative learning that may reduce the overall computational overhead is not considered.
[64]		CNN, DNN, LSTM	Resource allocation in SDN-based vehicular networks.	Reduce resource allocation time, enhance accuracy and throughput.	Gated Recurring Units and bidirectional LSTM could be considered to achieve better throughput.
[65]	RL	Q-learning	Resource allocation and uRLLC optimization for V2X networks.	Reduce the average end-to-end delay.	A distributed load-balancing strategy could be considered to improve networks' survivability and reliability.
[66]	DRL	DQN	Content caching and mode selection in UEs hotspot and UEs-V2I scenario.	Improve the learning speed and throughput, reduce the transmission delay.	The issues like robustness to model drift, handling outliers, and safe learning could be included in the future study.
[67]		DQN	Allocation of radio resources in vehicular network.	Enhance the data transmission success rate.	Radio and network resources could be allocated to downlink tasks to improve throughput.
[68]		DQN	Resource management and multi-vehicle task off-loading	Reduce computation complexity, improve the QoS of vehicular applications.	A greater number of vehicles causes the DQN output value space to increase, which reduces the system's performance.
[69]		DDPG	Blockchain-enabled Content caching in vehicular networks.	Achieve security and privacy protection in content caching,	Content transmission latency is not considered in V2V communication.
[70]		DQN	Autonomous resource provisioning in virtualized vehicular networks.	Balance resource utilization and QoS satisfaction level, improve overall system throughput.	The intra-cell interference management is not addressed in throughput maximization.
[71]		DQN	Transmission mode selection and resource management in V2X networks.	Maximize the sum capacity of V2I links, reduce latency in V2V pairs.	The impact of packet priority in QoS enhancement and communication link quality are not considered.
[72]	FL	F-MARL	Channel selection and power control for V2V communication.	Maximize the cellular sum rate and V2V packet delivery rate, reduce V2V transmission delay.	The adaptive switching between centralized and decentralized channel access for the V2V pair could be considered.

C. FOG RADIO ACCESS NETWORKS (F-RANS)

The Fog Radio Access Network (F-RAN) combines fog computing with radio access networks [21]. In F-RANs, computation and storage are brought closer to users at the network edge, reducing service latency and traffic load. F-RANs can be key enablers for the 6G wireless applications in industrial IoTs, D2D communications, and V2X communications due to their innate computing capability and storage capacity to overcome the limitations of custom C-RANs [32]. IoT devices generate large amounts of data, which places an enormous burden on traditional mobile networks [22]. Given this, the employment of AI models, specifically machine learning algorithms implemented in the F-RAN domain, enables AI-driven F-RAN. Besides providing real-time optimization for F-RAN, it improves delay performance, saves energy in content delivery, and reduces transmission burdens incurred by network data collection [41]. Recent studies for ML-assisted Fog-RANs are reviewed in this section. The presented work is classified based on the type of ML (SML, uSML, DL, RL, DRL, and FL) and the application scenario in vehicular networks as shown in Fig. 6.

Unsupervised Machine Learning (uSML): ML techniques have been implemented in F-RANs to address many optimization problems related to resource allocation, task

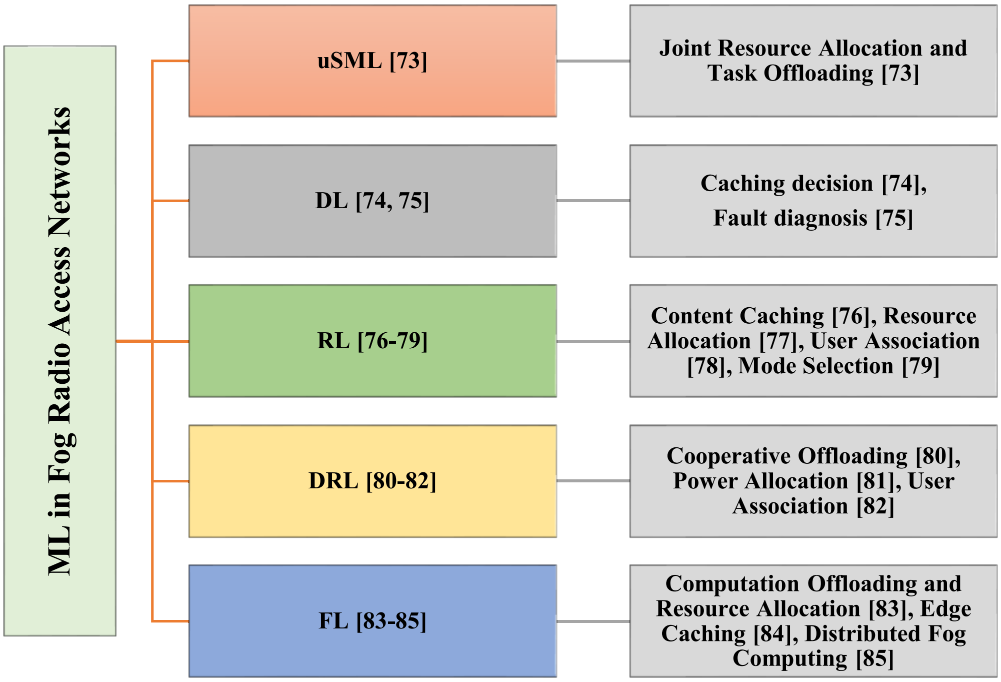


FIGURE 6. Recent work classification for ML techniques in fog radio access networks.

offloading, and computation-efficient communication. For instance, a two-stage resource-sharing and task-offloading mechanism is discussed in [73]. The authors developed a task-offloading algorithm using computational intelligence and contract theory to optimize the delay performance of each UE. Multi-armed bandits (MAB) online learning capabilities are utilized to alleviate the exploitation-exploration constraint in the online learning of task offloading. The authors proposed the DOT-VUCB method that allows UEs to learn the delay performance for each server based on the occurrence time and distance of fog servers. However, the proposed solution does not address cooperative communication between the fog servers.

Deep Learning (DL): A learning-based popularity prediction framework is proposed in [74] to address smart caching decisions in F-RANs using Bi-LSTM. The authors formulated a joint proactive-reactive caching policy using content replacement, user location prediction, and popularity prediction in F-RANs. The trend classes are determined using the k-Nearest Neighbor classifier and the popularity prediction models. It ensures the accuracy of the popularity prediction using deep learning, followed by the training of prediction models for each trend class. The proposed approach maximizes the cache hit ratio. However, the computation complexity of the proposed mechanism is not considered. In [75], a deep transfer learning-based mechanism is proposed for fault detection in F-RANs. It exploits the core-level information for the detection and labeling of the fault data based on spatial clustering. The authors exploited unsupervised deep transfer learning with a combination of CNN and a domain adversarial neural network for the classification of fault data. The proposed approach gives better convergence and fault detection accuracy. However, privacy and adaptability issues are not considered in the proposed methodology.

Reinforcement Learning (RL): In [76], the authors addressed caching applications in the F-RANs using Reinforcement Learning. A multi-arm bandit (MAB)-based AI-driven edge caching algorithm is proposed as a non-stationary version of the classic MAB algorithm considering the fog nodes and cloud as agents. This study addresses the challenges of traffic-demand variations and limited cache space at fog nodes considering traffic-demand variations, limited cache space at fog nodes under spatial-temporal variability, and unknown content popularity distribution. Compared to the benchmark schemes, the proposed caching scheme improves the cache space at the fog nodes and the cloud while reducing the average E2E delay. Nonetheless, the computation complexity of the proposed mechanism is not considered.

Multiple recent works [77], [78], [79] show that RL can effectively address resource management challenges n F-RANs. In [77], Q-learning with policy gradient is used for radio and computation resource allocation in Fog-assisted IoT networks. The authors used an actor-critic approach for real-time resource allocation while minimizing the task delay under the QoS constraints. However, this study does not consider the users' mobility, which affects the system's reliability. In [78], a Q-learning-based mechanism is proposed for the joint optimization of content placement and user association in F-RANs. It maximizes the network payoff and enhances prediction accuracy. The suggested scheme has been confirmed to minimize the overall computation complexity. However, the dimensionality and convergence issues for the large state and action space are not addressed.

In [79], the authors proposed a joint optimization framework for resource allocation and mode selection in uplink F-RANs. It aims to reduce the system power consumption and queue delay using the Q-learning-based DRL method. A joint optimization problem of mode selection and resource allocation is developed. Nonetheless, the issues like robustness improvement to model drift and signaling overhead are not considered. In [80], the authors investigated DRL for resource allocation and in-network caching optimization in a layered F-RAN. The proposed approach takes cross-layer cooperative caching and routing decisions using the DQN algorithm. It utilizes the systems' available network resources and the request history data of the incoming content requests. A neural network with a weighting parameter is used as an evaluation network in the proposed DQN. It also uses the optimal caching and routing options to minimize network delay. However, the proposed framework considers only the routing and in-network caching issues in HetNets.

The authors in [81] investigated power allocation and cache placement in F-RAN using deep Q-learning. F-APs are considered for cooperative service provisioning at the edge, while the DRL controller reduces the overall latency within the constraints of per-UE QoS requirements. At each decision step, it manages the allocation of the power resources and the cache placement decision. The proposed method gives higher convergence performance with lower latency than the baseline approaches. Nevertheless, the proposed model could be extended for distributed resource management in F-RANs. In [82], a DQN model is proposed to optimize the UE association for computation offloading in F-RAN. The proposed model consumes less energy than the random association and greedy association methods. However, this approach does not consider the collaboration of heterogeneous fog nodes.

Federated Learning (FL): Federated learning is an emerging technique for communication-efficient and privacy-aware Fog-assisted wireless networks [83], [84], [85]. Joint optimization of the power control and computation latency in F-RANs is discussed in [83]. The authors proposed a federated deep reinforcement learning (DRL) based algorithm to train the DDPG agents in each F-AP. It aims at minimizing the task execution delay and energy consumption of mobile devices while maintaining user privacy. However, the overall learning time could be further improved by exploiting gradient sparsification.

FL-assisted cooperative hierarchical (FLCH) edge caching framework for IoT networks is proposed in [84]. It enables smart caching decisions using cooperative hierarchical edge caching in Fog-RANs by training a shared global learning model at the F-AP. FLCH takes advantage of vertical and horizontal collaborations to optimize the cache hit ratio and latency of users. However, the convergence of the proposed method is not considered. In [85], task offloading using Federated Deep Q-Learning (FedDQL) in a vehicular fog computing (VFC) network is investigated. The authors considered the collaborative dew-enabled computation offloading mechanism to optimize the latency and computation cost in the VFC network while minimizing the energy consumption of MFC servers. However, the proposed approach does not consider collaborative communication between MFC servers.

Summary: As a summary, some of the recent studies using ML techniques in Fog-RANs are outlined. The proposed ML approach and an overview of the advantages and limitations of the recent works are provided in Table 5. In addition, the recent works are categorized according to the type of ML and the application scenario in Fog-RANs, such as joint resource allocation and task offloading [73], caching decision [74], fault diagnosis [75], content caching [76], resource allocation [77], user association [78], [82], mode selection [79], cooperative offloading [77], power allocation [81], computation offloading [83], edge caching [84], distributed fog computing [85] as shown in Fig. 6. These recent works focused to maximize the cache hit rate, optimize the system power consumption and reduce the queue delay and E2E delay.

Nonetheless, the current research has several limitations, including computation complexity, privacy, adaptability issues, and the users' mobility that could affect the system's reliability, which need to be addressed and considered further.

IV. MACHINE LEARNING TECHNIQUES IN RESOURCE MANAGEMENT TECHNICAL CHALLENGES

The upcoming 6G networks will support bandwidth-hungry applications including extended reality, the Internet of Senses, zero-touch cognitive networks, and autonomous driving. These emerging wireless network applications require the network services to be offered in such a way as to meet user experience requirements with various network performance characteristics, such as resource allocation, task offloading, mobility management, energy efficiency, and latency minimization. Therefore, there is a tremendous need to address these challenges using ML-empowered emerging wireless network architecture proposed in Fig. 3 to employ wireless networks' limited resources as effectively as possible. The recent work on machine learning-based methods for addressing the above-mentioned technical challenges is reviewed in this section.

A. RESOURCE ALLOCATION

In wireless networks, traditional resource allocation issues are modeled and solved as optimization problems. Due to diversified QoS and QoE requirements of upcoming 6G networks, resource optimization problems become mixed integer nonlinear programming (MINLP) NP-hard problems. Furthermore, channel conditions and user traffic parameters are challenging to obtain in such dynamically changing systems [86]. In recent works [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], ML techniques have been implemented in resource allocation problems, including

TABLE 5. Recent work summary of ML techniques in fog radio access networks.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/ Future work
[73]	uSML	DOT-VUCB	Joint optimization of resource allocation and task offloading in Fog-assisted IoT network.	Minimize the long-term delay.	The cooperative communication among the fog servers is not considered.
[74]	DL	Bi-LSTM	Smart caching decisions in F-RANs using popularity prediction models.	Maximize the cache hit rate.	The computation complexity of the proposed mechanism is not considered.
[75]		CNN, DANN	Fault detection in a fog radio access network.	Enhance the detection rate, improve accuracy, better convergence.	The privacy and adaptability issues are not considered.
[76]	RL	MAB	Caching application popularity for AI-driven Fog-RANs.	Improve caching space at the cloud and fog nodes, reduce average E2E delay.	The computation complexity of the proposed mechanism is not considered.
[77]		Actor-Critic	Joint optimization of radio and computation resource allocation.	Minimize the task delay while satisfying the QoS constraints.	The users' mobility that affects the system reliability is not considered in this study.
[78]		Q-learning	Joint optimization of content placement and user association in F-RANs.	Maximize the F-RAN network payoff, enhance accuracy prediction effect.	Large state and action spaces with high dimensionality and convergence are not considered.
[79]		Q-learning	Resource allocation and mode selection in uplink F-RANs.	Optimize the system power consumption and queue delay.	The issues like signaling overhead and robustness improvement to model drift could be considered.
[80]	DRL	DQN	Cloud-edge cooperative offloading in heterogeneous Fog-RAN.	Low-latency content transmission, Integrated RA to improve content distribution.	The proposed framework is limited to in-network caching and routing in HetNets.
[81]		DQN	Joint optimization of power allocation and cache placement in F-RAN.	Reduce latency while working under the limitations of per-UE QoS.	The proposed scheme could be extended for distributed resource management in F-RANs.
[82]		DQN	Computation offloading concerning user association in F-RAN.	Less system energy consumption in comparison with the random association and greedy association.	The collaboration of heterogeneous fog nodes could be addressed.
[83]	FL	F-DRL, F-DDPG	Computation offloading and resource allocation in F-RANs.	Minimize the task delay and energy consumption of mobile devices.	The overall learning time could further be improved by exploiting gradient sparsification.
[84]		F-DL	Smart caching decisions using cooperative hierarchical edge caching in F-RANs.	Reduce users' latency, enhance the cache hit ratio.	Blockchain-based content caching could be investigated to address data security.
[85]		F-DQL	Collaborative computation and offloading in vehicular Fog networks.	Optimize the computation cost and latency of task execution, reduce energy consumption of MFC servers.	The collaborative communication between MFC servers could be considered.

channel allocation, computation offloading, transmit power control, multi-constraint optimization concerning QoS, user association, and transmission mode selection. Recent studies for ML-assisted resource allocation are reviewed in this section. This recent work is classified into machine learning types (SML, uSML, DL, RL, DRL, FL) implemented in resource allocation, as shown in Fig. 7.

Supervised Machine Learning (SML): Resource allocation in wireless edge networks is proposed in [87]. A joint optimization of system energy consumption, end-to-end latency, and learning efficiency in Edge Machine Learning (EML) is formulated. In the proposed approach, two different RA schemes are considered. The first scheme considers the model-based cases, which contemplates energy consumption estimation using Least Mean Squares (LMS) under learning accuracy and latency constraints. The authors proposed a data-driven method for online performance evaluation of model-free cases. The second scheme maximizes the learning accuracy using SVM and NN classifiers for ML task inference at the edge server. The proposed approaches are quite flexible and adaptable to supervised, semi-supervised, and unsupervised learning applications. Nonetheless, the more sophisticated encoders, such as a vector quantizer could be used to achieve better rate-distortion pair.

In [88], the authors proposed deep learning for resource allocation in wireless edge networks. A CNN and SVM-based energy and time allocation framework is proposed. A statistical model based on supervised learning predicts the learning

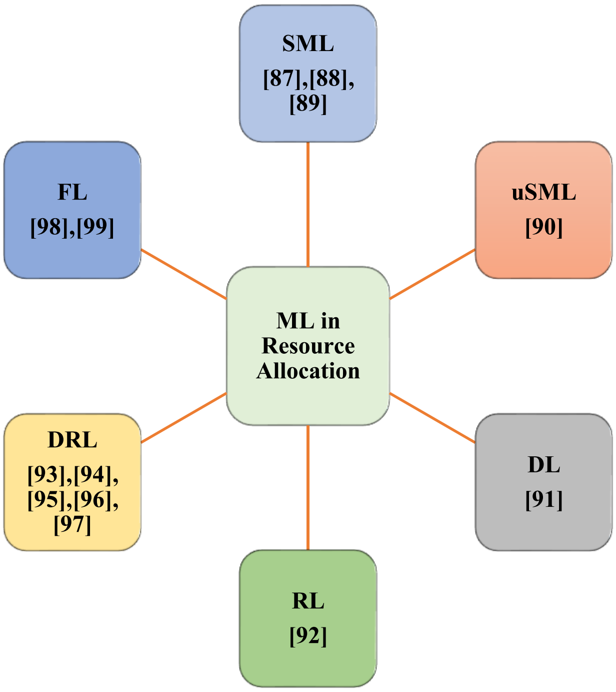


FIGURE 7. Recent Work Classification for ML Techniques in Resource Allocation.

accuracy across different tasks using the amount of available training data. The learning performance and wireless resources are analyzed using DCP and asymptotic solutions. It shows that the transmission time and generalization error are inversely proportional to each other concerning power allocation. The proposed algorithm is robust in practical and complex environments for resource allocation. However, it considers the statistical learning mechanics to predict learning accuracy.

In [89], the authors used transfer learning to address MINLPs for resource optimization in wireless networks. The proposed branch-and-bound algorithm learns the best pruning strategy to formulate a sequential decision problem. It exploits the sparse training examples of the problem data to achieve optimal performance and reduces the computation complexity. The mismatch issue is addressed by self-imitation, leading to an optimization mechanism based on transfer learning. This pruning policy uses a neural network-based classifier to optimize the computation complexity and performance efficiency. The suggested framework learns the optimization strategy with a few problem instances and is adaptable to small situation modifications. However, its performance deteriorates by changing the scenario parameters.

Unsupervised Machine Learning (uSML): In [90], an unsupervised ML-based framework is proposed to address subcarrier allocation in power domain-NOMA systems. The proposed scheme uses K-means clustering to form the clusters of the users. It facilities the subcarrier allocation based on associated channel properties to reduce the overall electromagnetic field exposure. However, the given approach does not consider multi-antenna systems.

Deep Learning (DL): The authors in [91] investigated a deep learning model for resource allocation in personalized wireless networks. A DNN-based surrogate model is proposed with a framework to manage surrogate models while aggregating real-time feedback on user satisfaction. It enhances the user satisfaction level by allocating the optimum resource blocks. Moreover, it gives operators more operational flexibility in terms of resource consumption rates and personalized user satisfaction. Nonetheless, personalized networks with surrogate assistance are not able to maintain the desired level of satisfaction.

Reinforcement Learning (RL): Reinforcement learning techniques are employed in resource management issues considering various application scenarios, including heterogeneous networks, ultra-dense networks, NOMA-based systems, and IoT applications. For example, the authors of [92] investigated the SBS control algorithm based on multi-agent Q-learning for resource allocation in ultra-dense wireless networks with small base stations. It minimizes the system energy consumption and reduces the number of outage users in uniform and non-uniform spatial distributions. Despite the rise in user mobility, the proposed method performs better than traditional algorithms like distributed Q-learning, random action, no transmit power control (TPC), and adaptive TPC. However, the agent only considers its state while calculating the computation complexity.

Deep Reinforcement Learning (DRL): In [93], a DRL-based intelligent resource distribution scheme under NOMA uplink transmissions is proposed. In three different traffic densities, the authors used DRL and SARSA algorithms to optimize the system sum rate for IoT users. Numerical results depict that the sum rate achieved in the proposed system is higher than the rate obtained in orthogonal multiple access systems, and it enhances the average rate. However, the convergence criteria, which may affect the sum rate, are not considered. In [94], an intelligent time division duplex (TDD) resource allocation scheme for the uplink and downlink HetNets is proposed. A DNN is developed that extracts the relevant features based on network information. The authors proposed a dynamic reinforcement learning model combining experienced replay memory with a dynamic Q-value iteration based on evaluated rewards. Although the network throughput and the packet loss rate have been considerably improved yet the computation complexity of the proposed DRL approach is not addressed.

A framework for resource management and intelligent network selection in Multi-RAT HetNets is discussed in [95]. The authors proposed Deep Multi-agent Reinforcement Learning (DMARL) to optimize the rewards in network selection and resource allocation for autonomous end-users and RANs, respectively. Each group of agents optimizes system performance simultaneously based on energy consumption, latency, and learning costs. Besides, in [96], MADRL is utilized for the state selection of small base stations (SBS).

It optimizes resource allocation and massive access in UDNs. The proposed approach enhances the system throughput and packets' successful transmission probability. However, the convergence of the MADRL algorithm could be improved by minimizing the end-to-end delay.

In [97], resource allocation among the tenants for service delivery is addressed to maximize the monetization of the infrastructure of future wireless networks. Based on DRL, an algorithm for autonomous learning of the optimal acceptance policy is proposed that satisfies the tenants' service guarantees. The proposed method gives optimal performance under a wide range of configurations. However, the performance is not validated using practical settings for resource management in the HetNet scenario.

Federated Learning (FL): In [98], the authors investigated joint optimization of task offloading and resource allocation in connected automated vehicular networks. A federated reinforcement learning (FRL) framework is proposed to reduce the execution delay of the optimal task offloading and computation resource allocation. Simulation results show that the proposed scheme minimizes the communication overhead and enhances the system throughput. However, the mobility of devices that may affect the system's power consumption is not considered. In [99], a federated-DDPG-based framework is presented to address joint optimization of power allocation to mobile users, phase shifts of mobile RIS, and RIS deployment in NOMA networks. The proposed method performs better than the benchmark schemes considering sum rate, channel quality, and training latency. Nevertheless, cooperative beam forming using multiple RISs could be considered.

Summary: Some recent studies using ML-assisted resource allocation in 6G networks are provided. The proposed ML approach and an overview of the advantages, and limitations of the recent works are listed in Table 6. Furthermore, a classification of recent works is provided based on ML types such as SML [87], [88], [89], uSML [90], DL [91], RL [92], DRL [93], [94], [95], [96], [97], and FL [98], [99] to address resource allocation problems as depicted in Fig. 7. These recent works focused on enhancing the system throughput, decreasing the packet loss rate and reducing the average load of base stations. However, several limitations to the current research, including network heterogeneity, transmit power constraints, and the convergence criteria that may affect the sum rate, need to be addressed and considered further.

B. TASK OFFLOADING

A significant number of latency-aware and computation-intensive tasks will be involved in 6G wireless networks due to multiple radio access technologies, slices, servers, and upcoming vehicular applications such as autonomous vehicles, vehicle-to-everything connectivity, and vehicular multimedia applications [100]. The ability to make real-time offloading decisions is a major challenge in reducing latency and energy consumption. Further, offloading decisions need to be made in a matter of milliseconds because channel conditions and other system variables are constantly changing. As the number of users and tasks increases, listing every possible decision can be challenging. Task offloading in Multi-access edge computing (MEC) systems is investigated in the recent literature [101], [102], [103]. Recent studies for ML-assisted task offloading are reviewed in this section. This recent work is classified into machine learning types (SML, uSML, DL, RL, DRL, and FL), addressing task offloading scenarios as shown in Fig. 8.

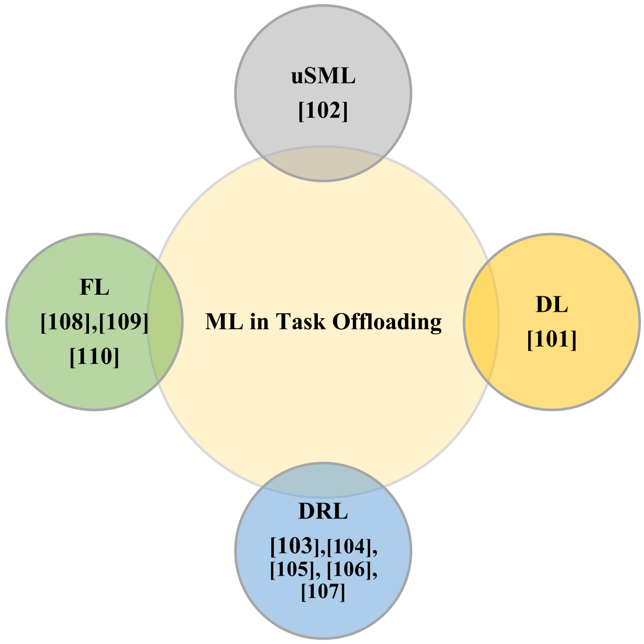


FIGURE 8. Recent work classification for ML techniques in task offloading.

Deep Learning (DL): Several recent works investigate machine learning-based dynamic computation offloading in wireless networks [98]. For example, in [101], a Distributed Deep learning-driven Task Offloading (DDTO) algorithm is presented to take offloading decisions in the heterogeneous network. It adaptively modifies parameters to make near-optimal offloading decisions by learning from past offloading events in MEC and MCC heterogeneous environments. As a result, it uses a high-dimensional search space to lower the cost of dimensionality. However, the proposed approach does not address the privacy and adaptability issues.

Unsupervised Machine Learning (uSML): In [102], the authors designed a task-offloading model by integrating blockchain technology and ML. The blockchain and smart contract-based mechanism addresses the privacy and fairness issues in secure task offloading. The authors used Merkle hash tree and smart contract to implement "proof-of-computing". Further, an online learning scheme for intelligent task offloading is developed based on QUeuing-delay aware, hand Over-cost aware, and Trustfulness Aware UCB (QUOTA-UCB) algorithm. The user vehicle learns the optimal long-term strategy based on task offloading delay

TABLE 6. Recent work summary of ML techniques in resource allocation.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/ Future work
[87]	SML	SVM, NN	Resource allocation in wireless edge network for ML tasks inference.	Enhance learning accuracy, reduce end-to-end service delay.	The more sophisticated encoders such as a vector quantizer could be used to enhance rate distortion pair.
[88]		SVM	Resource management based on energy and time allocation in wireless edge systems.	Maximize the learning performance, effective and robust in real systems.	The proposed model limits the predicting learning accuracy based on statistical mechanics of learning.
[89]		TL	Resource optimization in wireless networks.	Reduce the sample complexity, effectively handles task mismatch.	FL-based methods could be investigated to minimize computational complexity.
[90]	uSML	K-means clustering	Subcarrier allocation in power domain-NOMA environment.	Reduce the electromagnetic field exposure, optimize the transmitted power.	The proposed approach could be extended to multi-antenna systems.
[91]	DL	DNN	Optimization of resource allocation in personalized wireless networks.	Enhance user satisfaction level with optimum resource blocks.	Personalized networks with surrogate assistance could not maintain the desired level of satisfaction.
[92]	RL	Multi-agent Q-learning	Resource management in ultra-dense wireless networks with small base stations.	Improve energy efficiency, reduce the outage probability of users.	The agent only considers its state while calculating the computation complexity.
[93]	DRL	DRL, SARSA	Intelligent resource distribution under NOMA uplink transmissions.	Enhance the overall sum rates, achieve guaranteed average rates for IoT users.	The convergence criteria of the suggested method which may affect the sum rate, is not considered.
[94]		DQN	Dynamic allocation of radio resources using intelligent TDD configuration in HetNets.	Improve the network throughput, decrease the packet loss rate.	The computation complexity of the proposed DRL approach is not addressed.
[95]		MADDPG	Resource management and intelligent network selection in Multi-RAT HetNets.	Reduce energy consumption, improve latency while performing network selection.	The convergence of the MADDPG that may affect the end-to-end latency is not considered.
[96]		N3AC	Resource allocation among tenants for service delivery using network slicing.	Maximize the monetization of the infrastructure, improve the performance under different configurations.	The proposed approach could further be used in practical settings for resource management in the HetNet scenario.
[97]		DQN	Resource management in the Ultra-Dense NOMA System considering HTC and MTC.	Enhance the system throughput, reduce the average load of base stations.	The impact of information exchange between BSs could be considered in the proposed approach.
[98]	FL	F-RL	Resource allocation and offloading in vehicular networks.	Reduce the execution delay and communication overhead, enhance the system throughput.	The mobility of devices that may affect the system's power consumption is not considered.
[99]		FL-DDPG	Power allocation for users in mobile RIS-enabled NOMA Networks.	Optimize sum rate, enhance channel quality, reduce training latency.	The cooperative beam forming using multiple RISs could be considered.

minimization subject to long-term queuing delay and handover cost constraints. Compared to the conventional methods, this approach is robust and provides better performance regarding convergence, task offloading delays, and handover costs.

Deep Reinforcement Learning (DRL): DRL methods have been in trend in recent years to address optimization problems related to computation offloading due to the adaptability of DRL in unpredictable environments and complex tasks offloading scenarios. For example, in [103], a DRL-based computation offloading method in machine-type communication devices (MTCDs) is presented. MTC devices minimize the energy consumption in their uplink domain, considering delay-tolerant and non-delay-tolerant scenarios. It allows the MTCDs to achieve ideal offloading without being aware of MTC edges and cloud servers. Compared to cloud servers, deploying edge servers decreases the transmission distance, enhancing energy efficiency and QoS.

The authors in [104] proposed a DRL method for D2D offloading and content caching using an incentive mechanism. The estimated optimal solution is obtained using DQN, and a typical Vickrey-Clarke-Groves (VCG)-based payment rule is proposed to cover the cost of mobile nodes. The reverse auction mechanism maximizes the saving cost of the CSP and the offloading rate under different scenarios. In [105], a multi-agent double deep Q-learning (MA-DDQN) approach is presented for partial offloading and binary offloading in MEC-enabled wireless networks. It considers joint offloading and resource allocation in the uplink channels. In DDQN, different values are used while selecting and evaluating actions to mitigate over-optimistic estimation of the Q-value function. The proposed scheme enhances the computation efficiency and reduces the uplink power and the cost of computation energy.

The authors in [106] investigated DRL for computation offloading in a NOMA-based multi-access edge computing system. It accommodates the wireless channel variability and trains the DNN using the experience replay instead of Q values. It maximizes the weighted sum computation rate with binary computation offloading and gives higher computation rates than the OMA-based algorithm scheme. In [107], the authors used a MADRL-based scheme for joint task offloading and multi-channel access in MEC for industrial IoT networks. The machine-type agents (MTAs) cooperate for channel access and task offloading to reduce the computation delay. It formulates a policy using historical observations and actions without prior knowledge of the system parameters. The suggested method uses gradual learning to provide a low-complexity learning solution for multichannel access.

However, in most cases, DRL algorithms aggregate data to train centralized models. This centralized training mode may not be effective and raises the possibility of data leakage in light of the expanding complexity of mobile networks with increased flexibility of parameters. Additionally, implementing the DRL algorithms in large-scale MEC networks has several issues. For instance, as the number of edge devices increases, the state-action space of the MEC system is exponentially increased, which is in-efficient for Q-learning and DQN algorithms. In addition, the multi-agent DRL (MA-DRL) algorithm fails to perceive the global environment resulting in converging into the optimal local solution.

Federated Learning (FL): Federated learning has the advantage of privacy protection of personal data and has been recently employed in task offloading problems in MEC systems [108], [109], [110]. Data privacy is maintained by FL, as it transfers parameter updates instead of the original data to the server. The authors in [108] proposed an online algorithm using a combination of FL and DRL in a wireless-powered communication-MEC system for task offloading and resource allocation. In the proposed system, a base station executes computation-intensive tasks exploiting task offloading. A learning rate adjustment is investigated to improve FL convergence with non-IID data. The proposed method gives better performance in terms of convergence time, CPU execution delay, and stability as compared to traditional numerical optimization methods. However, the dynamics of tasks under system environment constraints are not addressed in the proposed approach.

In [109], the problem of computation offloading and service caching placement in UDNs is addressed. The authors proposed FL to train the two-timescale deep reinforcement learning to protect the personal data privacy of edge devices. A hybrid computation offloading strategy enables edge users to offload resource-intensive tasks to the remote cloud servers, nearby edge servers, or nearby mobile devices using D2D computation offloading. The suggested method reduces the overall offloading time and network resource consumption by coordinating the optimization of compute offloading, resource allocation, and service caching placement. However, this study does not consider the users' mobility that affects the spectrum usage. In [110], the authors investigated computation offloading in space-assisted vehicular networks. An asynchronous federated DQN-based algorithm is proposed to address the task offloading in user vehicles. The suggested method reduces the system delay and maximizes the system's throughput while maintaining the system's reliability. However, the cost of computation energy is not addressed in the proposed approach.

Summary: Some recent works addressing task offloading using ML techniques in 6G networks are discussed. The proposed ML approach and an overview of the advantages and limitations of the recent works are provided in Table 7. Furthermore, the recent works are categorized according to the ML types such as DL [101], uSML [102], DRL [103], [104], [105], [106], [107], and FL [108], [109], [110] to address task offloading problems as explained in Fig. 8. These recent works focused to increase the computation efficiency, optimize resource utilization, enhance the convergence speed, and reduce the system delay. However, there are several limitations to the current research, including the users' mobility that may affect the spectrum usage, the impact of intra-cell interference, and collaborative edge caching in heterogeneous environments, which need more attention and consideration.

C. ENERGY EFFICIENCY

Energy efficiency has a significant importance in designing 6G devices due to their utilization of higher frequency bands. Intelligent and self-healing machines, including autonomous vehicles and connected drones, require communication links in D2D communication and connectivity with base stations which increases the concern of energy consumption about sustainable wireless infrastructure. Further, in MEC-enabled IoT networks, fog nodes perform low-latency computation offloading with limited storage and power resources. The traditional algorithms for supervised learning, reinforcement learning, and modern federated learning techniques can reduce energy consumption and enhance communication efficiency in 6G networks. A low-power communication system integrated with an energy-efficient computing mechanism could help enhance QoS and QoE for end users in addition to providing a low-cost and sustainable communication infrastructure for 6G wireless networks. Recent studies for ML-based energy-efficient communications are reviewed in this section. This recent work is classified into machine learning types (SML, uSML, DL, RL, DRL, FL) implemented to enhance energy efficiency in wireless networks, as shown in Fig. 9.

Supervised Machine Learning (SML): Several recent works employed machine learning techniques to address energy efficiency in various next-generation wireless applications. For instance, the authors in [111] used supervised machine learning for cache localization in D2D communications. It reduces the access delay of UEs and minimizes energy consumption while predicting accurate cache

TABLE 7. Recent work summary of ML techniques in task offloading.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/Future work
[101]	DL	DNN, DDTO	Task offloading in MEC system with collaborative heterogeneous clouds.	Reduce the computation complexity.	Blockchain and meta-learning could be used to address privacy and adaptability issues.
[102]	uSML	QUOTA-UCB	Task offloading in fog-assisted vehicular networks.	Minimize the queuing delay, optimize the handover cost.	The proposed approach could be investigated for integrated vehicular networks in the space-air-ground scenario.
[103]	DRL	DQN	The computation offloading in machine-type communication devices.	Minimize the system power consumption, enhance performance gain of 12% in computation offloading.	The adaptability of the proposed algorithm in determining the DNN settings is not considered.
[104]		DQN	Content caching and task offloading in D2D communication.	Enhance the saving cost, cache hit ratio, and offloading rate.	Collaborative edge caching in heterogeneous environments is not considered.
[105]		DDQN	Learning-based partial and binary offloading in MEC networks.	Increase the computation efficiency, minimize computation energy.	Task dependencies for collaborative MEC servers are not considered.
[106]		DQN	Intelligent offloading in a multi-access edge computing system.	Optimize the computation rates of MEC systems.	The impact of intra-cell interference is not considered in proposed approach.
[107]		MADRL	Computation offloading for Industrial IoT users using MEC.	Minimize the computation delay, enhance the success rate of channel access.	Gated recurrent units could be included for better prediction of the channel conditions.
[108]	FL	F-DRL	Online task offloading in wireless-powered communication systems.	Minimize the CPU execution delay, enhance the convergence speed.	The proposed model is not validated for task dynamics concerning the overall energy efficiency of the system.
[109]		F-DRL	Service caching with computation offloading in ultra-dense networks.	Optimize resource utilization, reduce the task execution time.	The users' mobility that affects the spectrum usage is not considered.
[110]		F-DRL, F-DQN	Computation offloading in space-assisted vehicular networks.	Maximize the system's throughput, reduce latency while maintaining reliability.	The cost of computation energy is not addressed in the proposed approach.

placement locations using decision trees and random forests. A trust factor is used to classify nodes as compact, overlapping, and distributed. The authors conclude that the random forests algorithm gives the highest accuracy with trust compared to decision trees. Furthermore, in D2D communication, energy consumption is minimum; however, the energy consumption increases in content sharing between users and the gateways.

Deep Learning (DL): DL is a promising method to address non-convex optimization problems. Resource allocation in underlaying D2D communication exploiting DL has been investigated in [112]. The authors proposed the Dinkelbach algorithm with resource allocation to mitigate the cumulative interference in multi-D2D pairs. It allocates the multi-D2D pairs to one RB to address the global energy efficiency problem while both cellular users and D2D pairs fulfill the data rates requirement. The proposed approach shows less complexity compared to the exhaustive search and could be extended for ultra-dense heterogeneous environments to enhance the GEE.

Reinforcement Learning (RL): RL methods are employed to solve decision-making problems for energy-efficient wireless communication systems, such as end-to-end communication, user association, and resource management [113], [114], [115], [116]. In [113], RL is used to reduce transmission energy in wireless networks. The authors proposed MEC for small cells to offload traffic from the MBS, and small-cell users are introduced to cache-enabled D2D communication. Q-learning algorithm is applied to obtain an optimal caching policy, while the DQN algorithm is implemented for SBS.

Deep Reinforcement Learning (DRL): In [114], DRL based power optimization method in an underlay D2D communication network is proposed. The authors considered two parallel DQNs to enhance energy efficiency with guaranteed QoS considering the system throughput. BS selects DQNs based on the current system state for joint optimization of system throughput and EE. Nonetheless, with more users, the extent of the action space expands exponentially, requiring substantial exploration throughout training.

In [115], joint optimization of resource allocation and mode selection in a D2D-enabled heterogeneous network is discussed to maximize the long-term energy efficiency of D2D links. The user can select a cellular or D2D mode using a Markov decision process. The authors proposed a DDPG

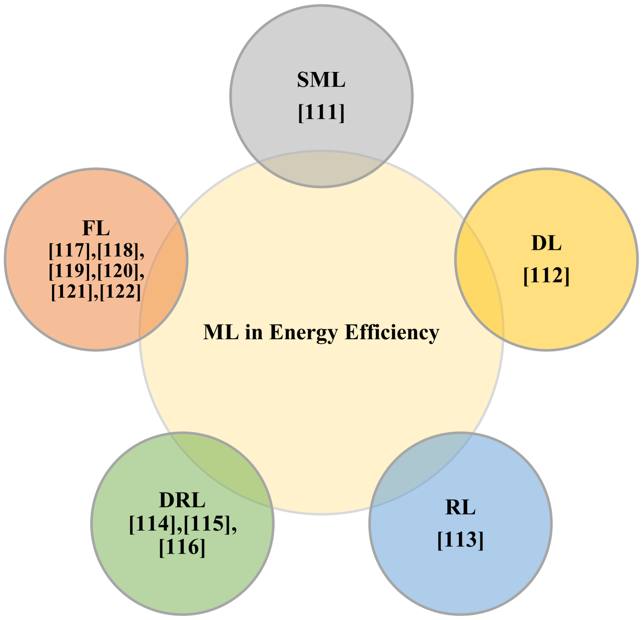


FIGURE 9. Recent work classification for ML techniques in energy efficiency.

(deep deterministic policy gradient) algorithm to achieve an optimal policy based on continuous state and action space. It enhances the convergence rate and gives better energy efficiency compared to other benchmark schemes. However, this study does not consider the interference incurred during D2D communications that could impact the QoS.

The authors in [116] considered a DRL model for resource allocation and task offloading in MEC. It aims at reducing the energy consumption of each user, considering the latency and limited computation resources. The authors proposed the multi-agent deep deterministic policy gradient (MADDPG) to solve the joint optimization problem. It performs centralized training and decentralized execution due to which it uses only the actor network for resource allocation in the testing phase. Further, MADDPG gives better efficiency and stability in dealing with the dynamic environment. The proposed algorithm improves convergence and reduces the users' energy consumption. Nonetheless, the computation complexity of the suggested method is not considered.

Federated Learning (FL): FL with collaborative learning capabilities has recently received attention due to the benefits of improved data privacy and energy efficiency [117], [118], [119], [120], [121], [122]. In [117], the authors proposed a computation task offloading mechanism in edge computing-enabled space–air–ground integrated networks. It enables the IoT devices to take real-time offloading decisions using the Federated DRL approach. An optimization problem is formulated to minimize the energy consumption of the computation tasks and the adaptive federated DRL algorithm considers the privacy protection and communication failure. The proposed method gives optimal offloading decisions with less energy consumption compared to the benchmark schemes. However, the computation complexity considering the user's mobility is not addressed.

Joint optimization of resource allocation and task offloading in IoT networks is discussed in [118]. In the proposed approach, federated-double deep Q-network (F-DDQN) learning is exploited to optimize transmit power allocation, offloading decisions, and computation resource allocation. The proposed approach shows better performance in terms of the task execution delay and energy consumption of IoT devices than benchmark schemes. Nevertheless, the computation complexity of the proposed approach is not addressed.

In [119], a federated-deep reinforcement learning-based framework is presented to address joint optimization of user association and power control in UAV-assisted multi-access edge computing systems. The proposed method performs better than the benchmark schemes considering the sum power, data rate, and computation cost. Nevertheless, a distributed load-balancing strategy could be considered to improve the system's throughput. In [120], the authors considered resource allocation and user association in high-altitude balloon networks. It cooperatively trains an optimal SVM model using users' data without prior knowledge of user association results. The proposed approach reduces the transmission energy and enhances the computation efficiency. However, other DL models could be considered with FL for predicting the optimal user association to reduce energy consumption.

The authors in [121] proposed distributed federated learning (DBFL) to address energy efficiency in distant IoT or edge devices. The proposed DBFL effectively handles heterogeneous devices and identifies device types among Edge, IoT, or vehicular devices using variable transmission delay. Experiment results show that the given framework enhances the classification accuracy and minimizes the system energy consumption. In [122], the authors addressed the energy efficiency in reconfigurable intelligent surface (RIS)-based indoor multi-robot communication systems. A Federated-DRL-based approach is proposed to optimize the transmit power of APs, robot trajectory, and phase shift of RIS. The proposed method reduces the network energy consumption and enhances model accuracy and learning efficiency. However, the transmit power analysis under the different scales of the networks could be considered.

Summary: Some of the recent studies addressing ML-assisted energy-efficient communication networks in 6G are discussed. The proposed ML approach and an overview of the advantages and limitations of the recent works are provided in Table 8. Besides, the recent works are categorized based on ML types such as SML [111], DL [112], RL [113], DRL [114], [115], [116], and FL [117], [118], [119], [120], [121], [122] to address energy efficiency problems as shown in Fig. 9. These recent works focused at maximize energy efficiency with guaranteed QoS considering system throughput, give better convergence, reduce the transmission energy, and reduce the training loss. However, there are several limitations to the current research, including the computation complexity considering the user's mobility and transmit

TABLE 8. Recent work summary of ML techniques in energy efficiency.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/ Future work
[111]	SML	DT, RF	Caching localization in heterogeneous IoT using D2D communication.	Minimize energy consumption during communication between the neighboring users.	Blockchain-based content fetching could extend the proposed model for future work.
[112]	DL	ANN	Resource allocation in underlaying D2D communications systems.	Enhance global energy efficiency.	The proposed work could further be exploited to increase GEE for multi-user, multi-cell systems.
[113]	RL	Q-learning	Caching selection in integrated D2D communication and cache-enabled IoT.	Minimize the cost of traffic transmission energy.	The computation complexity considering the user's mobility could further be investigated.
[114]		DQN	Dynamic power optimization in D2D pairs and cellular users.	Maximize energy efficiency with guaranteed QoS considering system throughput.	With an increase in the number of users, the action space exponentially increases which reduces the system's performance.
[115]	DRL	DDPG	Resource allocation and mode selection in a D2D heterogeneous network.	Better convergence, increase energy efficiency as compared to other benchmark schemes.	Interference incurred during D2D communications is not considered.
[116]		MADDPG	Resource allocation and task offloading scheme in MEC.	Reduce the energy consumption of each user.	The computation complexity of the proposed approach is not addressed.
[117]		F-DRL	Edge computing for the Internet of Remote Things.	Minimize the energy consumption of the tasks computations.	The computation complexity considering the user's mobility could be investigated.
[118]		F-DDQN	Resource allocation and task offloading IoT networks.	Minimize energy consumption of IoT devices, reduce the task execution delay.	The computation complexity of the proposed approach is not addressed.
[119]		F-DRL	Resource allocation in UAV-assisted multi-access edge computing systems.	Minimize the sum power, enhance data rate, reduce latency and computation cost.	A distributed load-balancing strategy could be considered to improve the system's throughput.
[120]	FL	F-SML, SVM	Resource allocation and user association in high altitude balloon network.	Reduce the transmission energy, enhance computation efficiency	Other DL-based models could be considered to predict the optimal user association.
[121]		DBFL	Edge, IoT, or vehicular device identification using variable transmission delay.	Enhance classification performance, reduce energy consumption in comparison to CVFL.	Future studies could include more robust verification and validation measures of DBFL.
[122]		F-DRL	Trajectory design for RIS-assisted indoor multi-robot communication system.	Reduce the network energy consumption, 86% less training time, more robust compared to centralized DRL.	Transmit power analysis under the different scales of the networks could be considered.

power analysis under the different scales of the networks, which need further investigations and consideration [123], [124].

D. LATENCY MINIMIZATION

The upcoming 6G communication systems are designed to operate at terabit-per-second data rates with extremely low latency. Massive MTCs and URLLCs have revolutionized cellular communications recently. MTC enables a wide range of intelligent IoT connectivity, including autonomous transport systems, zero-touch cognitive networks, and crowd sensing, to name a few. All these MTC scenarios have different QoS requirements. Further, a remote data center, i.e. a cloud server may not address the low latency constraints for content access in such applications. These challenges need to be addressed by integrating advanced ML techniques with 6G mobile network architectures and intelligent optimization techniques. Recent studies for ML-based latency minimization are reviewed in this section. This recent work is classified into machine learning types (SML, DL, DRL, FL) implemented in latency minimization in wireless networks, as shown in Fig. 10.

Supervised Machine Learning (SML): ML is a valuable tool for analyzing big data and making data-driven decisions to improve network performance by URLLC QoS standards. For instance, QoS optimization in IoT networks is addressed in [125]. The authors proposed an extended Kalman filtering (EKF) method for harvesting power prediction in IoT applications. The UEs then select security suites that meet their requirements while maintaining continuous service. The proposed approach achieves satisfactory security protection for various services and enhances efficiency and throughput by circumventing energy exhaustion.

Deep Learning (DL): The authors in [126] proposed a Fast uplink grant (FUG) resource allocation method for massive IoT. It prioritizes the machine-type communication (MTC) devices using an SVM classifier while LSTM is proposed for real-time MTD traffic prediction. The proposed FUG approach minimizes the access delay and enhances the system throughput. However, the learning-based selection of the

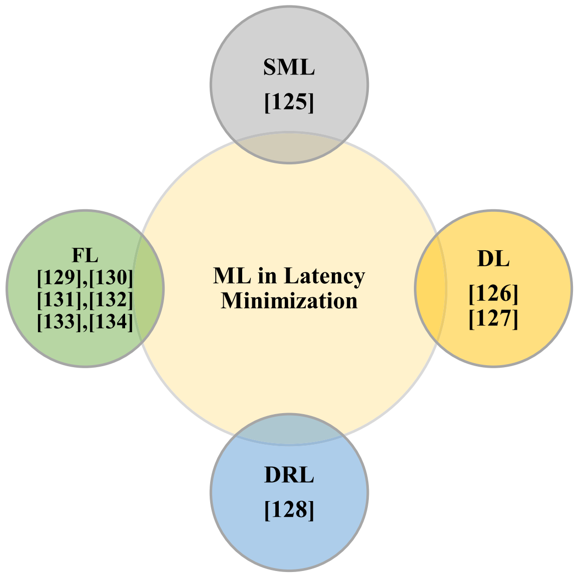


FIGURE 10. Recent work classification for ML techniques in latency minimization.

appropriate exploration could be further investigated. The authors in [127] addressed resource allocation using data sharing and distributed training in mobile edge networks. The distributed Batch Gradient Descent (BGD) is implemented for CNN training. The proposed approach minimizes the training latency and enhances the training speed and accuracy.

Deep Reinforcement Learning (DRL): In [128], DRL-based multi-channel access in high-mobility communication systems is proposed. The authors introduced deep deterministic policy gradients (P-DDPGs) to address the challenges of the high-dimensional action space and low convergence speed of DRL. The P-DDPG algorithm uses a learning-based DMCA framework that includes the channel prediction module (CPM) and the P-DDPG module. With this framework, the processing delay is effectively reduced by exploiting the features of the high-mobility system. It reduces the non-instant decision error with a faster convergence rate while making channel access policy decisions at each time slot. However, the E2E latency that could affect the system's reliability is not considered.

Federated Learning (FL): FL-based performance optimization of wireless networks in terms of latency minimization has been addressed in recent studies [129], [130], [131], [132], [133], [134]. The authors in [129] proposed federated learning to address node selection and cache replacement for collaborative edge caching in D2D-assisted HetNets. The proposed approach effectively reduces average delay and improves hit rate and average reward. However, the issues like security and privacy protection in cache replacement are not addressed. A federated deep reinforcement learning model is proposed in [130] for QoS enhancement in UAV-based vehicular networks. The proposed approach aims to minimize latency and increase communication reliability. Nonetheless, the system performance could be improved by considering fading effects in the proposed model.

In [131], decentralized cooperative edge caching in IoT networks is investigated. The authors proposed a federated DRL-based framework to address the duplicate traffic offloading and enable IoT devices for decentralized collaborative training. The proposed method reduces the average delay, enhances communication efficiency in terms of cache hit rate, and reduces performance loss. However, this technique is limited to performance improvement in the content access delay for UEs. The authors in [132] utilized the network slicing approach for QoS enhancement in Industrial-IoT (IIoT) applications. A deep federated Q learning (DFQL) method is proposed to train a global model to maximize the slices' rewards by leveraging agents' self-experiences. The proposed approach optimizes the self-QoS by enhancing throughput and reducing latency. However, the computation complexity of the proposed framework is not considered.

In [133], channel access optimization for new radio in unlicensed spectrum (NR-U) is investigated. The authors proposed a federated DRL to optimize the energy detection thresholds. It ensures a reliable packet delivery for downlink URLLC transmission in NR-U. The proposed approach maximizes communication reliability and reduces latency. However, the constraints on system fairness like proportionate fairness are not addressed in the proposed method. The authors in [134] proposed a federated DRL approach to address cooperative caching in mobile edge networks. It provides a minimal signaling overhead mechanism to share the model parameters with intelligent edge devices.

The proposed method trains a model in less time, provides better convergence, and enhances the overall cache hit ratio. However, the mobility of the users that may impact communication efficiency is not considered.

Summary: Some of the recent works addressing latency minimization using ML techniques in 6G networks are discussed. The proposed ML approach, and an overview of the advantages, and limitations of the recent works are provided in Table 9. Furthermore, the classification of the recent works is also given based on ML types such as SML [125], DL [126], [127], DRL [128], and FL [129], [130], [131], [132], [133], [134] to address latency minimization problems as shown in Fig. 10. These recent works focused on reducing the access delay, minimizing the processing delay, the stringent delay, and reduce the training latency with joint routing and spectrum allocation. However, several limitations to the current research, including the issues like security and privacy protection, and the E2E latency which could affect the system reliability need to be addressed and considered further.

E. HANDOVER MANAGEMENT

In 6G communication networks, handover management has a prime significance in maintaining QoS due to several challenges, such as reducing throughput and disrupting service. Furthermore, 6G networks will deploy more mmWave base stations so that all mobile terminals have a line-of-sight

TABLE 9. Recent work summary of ML techniques in latency minimization.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/ Future work
[125]	SML	EKF	QoS optimization in IoT networks.	Improve the throughput and working time, reduce the energy exhaustion.	The DL-based solutions for addressing the QoE could be investigated in future work.
[126]	DL, SML	LSTM, SVM	Fast uplink grant allocation in mMTC and IoT networks.	Reduce access delay, enhance prediction accuracy and throughput.	The learning-based selection of the appropriate exploration rate could be investigated.
[127]	DL	CNN	Resource allocation using data sharing and distributed training in mobile edge networks.	Minimize the training latency, improve the training time and the training accuracy.	The proposed approach is not considered for the IID data distribution at edge devices.
[128]	DRL	P-DDPG	Multi-Channel Access in high-mobility communication systems.	Minimize the processing delay, reduce the non-instant decision error.	The E2E latency which could affect the system reliability is not considered.
[129]	FL	F-DRL	Node selection and cache replacement for collaborative edge caching.	Reduce average delay, improve the hit rate, better performance in terms of average reward.	The issues like security and privacy protection in cache replacement are not addressed.
[130]	FL	F-DRL	Resource management in UAV-aided vehicular networks.	Enhance QoS in terms of reliable connectivity and minimize latency.	The system performance could be improved by considering fading effects in the proposed model.
[131]	FL	F-DRL, F-DDQN	Cooperative edge caching in IoT networks.	Minimize the average delay, improve the cache hit rate, reduce the performance loss.	The suggested technique is limited to performance improvement in the content access delay for UEs.
[132]	FL	DF-QL	Network slicing for Industrial IoT applications.	Maximize QoS by enhancing throughput and reducing latency.	The computation complexity of the proposed framework is not considered.
[133]	FL	F-DRL	Channel access optimization for new radio in unlicensed spectrum.	Maximize the reliability and reduce latency in downlink URLLC.	Constraints on system fairness like proportionate fairness are not addressed.
[134]	FL	F-DRL	Cooperative caching in mobile edge networks.	Minimize training latency, better convergence, enhance communication efficiency.	The mobility of the users that may impact communication efficiency is not considered.

(LOS) path to the deployed base stations. Thus, an important research direction is to develop ML-based decision solutions that identify the base station to initiate the handover. Recent studies for ML-based handover management are reviewed in this section. This recent work is classified into machine learning types (DL, RL, DRL, FL) implemented in handover management in wireless networks, as shown in Fig. 11.

Deep Learning (DL): In recent years, ML methods have been investigated for intelligent handover management in dynamic wireless environments to enhance communication efficiency. For instance, in [135] a deep learning model is proposed to manage handover decisions in B5G RANs. The authors used LSTM to select the next-Generation Node B (gNB) based on download time using previous experience instead of adjusting typical HO parameters. The proposed approach enhances the initial throughput by selecting a better MCS. Consequently, it improves the users' QoE. However, the impact of the HO process interrupts, which could affect the system throughput, is not considered.

Reinforcement Learning (RL): RL methods implemented by some researchers to optimize the handover parameters [136], [137], [138], [139]. The authors in [136] and [137] discussed Q-learning-based handover parameters optimization for mobility management of dynamic small-cell networks and reliable connectivity in mmWave networks, respectively. The proposed solutions reduce the adaptation time and improve the user satisfaction rate. Nevertheless, the state and action space expands considerably with more users, making the traditional RL inept for large-scale networks.

In [138], a multi-agent RL (MARL) is investigated for smart handover management considering the users' QoS requirements. It optimizes the handover efficiency, cost of handover, and outage probability. Nevertheless, the proposed approach could be extended to address power consumption and interference management. MARL addresses the Q-learning scalability and value function accuracy for data collection of each network user. The authors in [139] proposed MARL to address user association in dense multiple-radio access networks. The proposed mechanism limits the HO signaling overhead, achieves significant sum-rate gain, and reduces computation complexity. Nonetheless, SBS connectivity is not considered in the decision process.

Deep Reinforcement Learning (DRL): DRL-based algorithms use deep structures with a combination of reinforcement

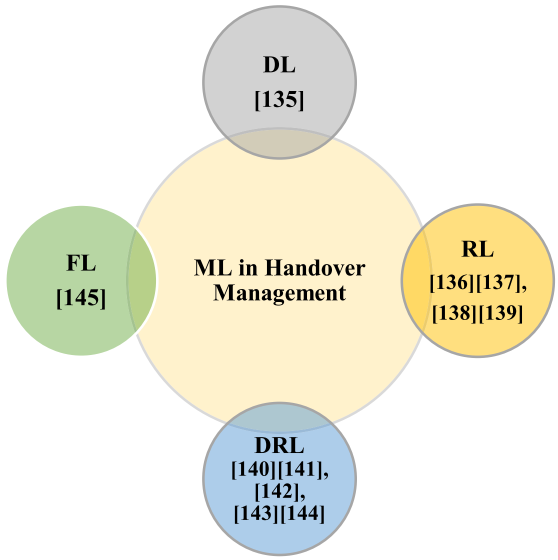


FIGURE 11. Recent work classification for ML techniques in handover management.

learning to understand the dynamic and changing environment using state estimation and function approximation. It derives the optimal decisions for long-term operations in various network applications [140]. In [141], DRL is applied for autonomous handover control considering user connectivity and throughput in mmWave communication. A DQN model intelligently takes action using an epsilon-greedy policy to enhance the throughput while decreasing the HO occurrence. Nonetheless, the transfer learning-based mechanism could further be investigated in an ultra-dense network environment. In another work [142], DDQN-based intelligent handover management is investigated for vehicle-to-network communications. It significantly reduces the HO delay and optimizes the packet loss compared to baseline methods. However, the inter-slice handover scenario in the network slice is not considered.

DRL-based joint optimization of power allocation and handover management in a heterogeneous environment is investigated in [143]. A multi-agent PPO (MAPPO) method is presented for the centralized training of multiple users with decentralized execution. The proposed MAPPO approach reduces the HO frequency and maximizes the overall throughput compared to the MADDPG algorithm. Nevertheless, ICI effects on system throughput are neglected in the proposed scheme. The authors in [144] used a model-free DRL method for adaptive switching between horizontal and vertical HOs for V2V and V2I communication. The proposed approach supports hybrid HOs and UC clustering to improve the connectivity-HO tradeoff. However, the communication and computation (dual functionality) of RSUs and APs are not considered.

Federated Learning (FL): A federated deep reinforcement learning method is presented in [145] for user access control in O-RANs. In the proposed approach, each UE makes access decisions with its DQN independently, while a global model server installed with the RAN intelligent controller (RIC) is responsible to update DQN parameters received from every UE for user access control. The proposed approach reduces the HO frequency and maximizes the long-term throughput compared

to the baseline methods. Nevertheless, the proposed scheme could be extended for multilayer networks with different control cycles.

Summary: Some of the recent studies addressing handover management issues exploiting ML methods are presented. The proposed ML approach and an overview of the advantages and limitations of the recent works are provided in Table 10. Furthermore, the recent works are categorized based on ML types, such as DL [135], RL [136], [137], [138], [139], and DRL [140], [141], [142], [143], [144], FL [145] to address handover management problems as depicted in Fig. 11. These aforementioned works focused on reducing handover frequency, improving handover efficiency, reducing the cost of handover, and minimizing the outage probability. However, several limitations to the current research, including the HO signaling overhead, the inter-slice handover scenario in the network slice, and the impact of the HO process interrupts affecting the system throughput need to be addressed and considered further.

F. LESSON LEARNED

In this section, we outline the major lessons learned from a comprehensive review of recent literature on the integration of ML techniques with advanced technologies in 6G networks. The following lists the lessons captured:

- ML-assisted D2D communication has the advantages of latency minimization and higher throughput efficiency for next-generation wireless networks. Furthermore, the proximity in D2D connections and reusability gain in D2D-enabled wireless communication systems make them energy efficient and reduce the congestion in cellular networks [146]. Nevertheless, interference management is challenging in such networks due to the absence of base stations [45]. The performance of D2D users can be adversely affected by high levels of interference from other D2D links and cellular users.
- The combination of ML techniques and the Internet of Vehicles (IoV) gives various opportunities for smart communications such as context-aware decision-making in autonomous driving, resource allocation, smart mobility management, task offloading, and load distribution for optimal deployment and path planning in V2X networks [59]. It also exhibits significant ITS enhancements, which would effectively be a transportation system for the coming years. Utilizing AI to drive intelligent transportation systems, highways, railways, aviation, and maritime courses can be made safer, more efficient, and environment-friendly. This supports economic growth and environ-

TABLE 10. Recent work summary of ML techniques in handover management.

Ref	Learning Type	Approach	Application Scenario	Advantages	Limitations/Future work
[135]	DL	LSTM	Handover decision to select a target gNB in B5G RANs.	Enhance initial throughput by selecting a better MCS, Improve users' QoE.	The impact of the HO process interrupts, which could affect the system throughput, is not considered.
[136]	RL	Q-learning	Handover parameters optimization for mobility management of dynamic small-cell networks.	Reduce adaptation time, improve the user satisfaction rate.	HO signaling overhead is not addressed in the proposed mechanism.
[137]	RL	Q-learning	Handover management for reliable connectivity in mobile mmWave networks.	Provide constant rate and reliability in the user's trajectory with minimum HOs.	The HO cost, which could impact the system throughput is not addressed.
[138]	RL	MARL	Smart handover for RAN slicing with diverse users' QoS requirements.	Improve handover efficiency, reduce the cost of handover, and minimize outage probability.	The proposed approach could be extended to address the power consumption and interference management.
[139]	RL	MARL	User association in dense multiple-radio access networks.	Limit the HO signaling overhead, provide large sum-rate gain, reduce computation complexity.	SBS connectivity is not considered in the decision process.
[140]	DRL	DQN	Handover timings optimization and predicting the data rate degradations in mmWave networks.	Better performance without the state space expansion, Low computational complexity.	More features with state information could be considered for predicting data rate degradation in HetNets and UDNs.
[141]	DRL	DQN	Autonomous HO control considering user connectivity and throughput in mm-wave communication.	Decrease the number of HO occurrences by 70% to the traditional approach, Enhance system throughput.	The transfer learning-based mechanism could be investigated in an ultra-dense network environment.
[142]	DRL	DDQN	Intelligent handover management in vehicle-to-network communications.	HO delay reduction by 11.56 s per HO, 25.73% packet loss minimization as compared to baseline methods.	The inter-slice handover scenario in the network slice is not considered.
[143]	DRL	MAPPO	Power allocation and HO management in heterogeneous networks.	Reduce the HO frequency, maximize the overall throughput.	ICI effects on system throughput are neglected.
[144]	DRL	DDPG	Adaptive switching between horizontal and vertical HOs for V2V and V2I communication.	Support hybrid HOs, improve the connectivity-HO tradeoff by 30%.	The computation complexity for both RSUs and APs is not addressed.
[145]	FL	F-DRL	User access control in O-RANs.	Reduce the HO frequency, maximize the long-term throughput	The proposed scheme could be extended for multilayer networks with different control cycles.

mental sustainability. Moreover, ML techniques have been a promising candidate to address these challenges in dynamic environments such as vehicular networks.

- F-RANs can relieve traffic congestion by using ML-enabled edge caching and edge computation, which improves delay performance and conserves energy. F-RANs have the potential to facilitate the IoV, context-aware resource allocation, and latency-aware connectivity to commuters and integrate intelligent parking and smart traffic lights. Fog nodes need to have varying characteristics depending on the IoV application. A wide range of industries with diverse requirements, including consumer, wearable, industrial, and enterprise goods, automobiles, and healthcare, are starting to embrace the Tactile Internet, Industrial Internet of Things (IIoT), Internet of Everything (IoE), and Internet of Senses (IoS). Therefore, a large amount of data will be used for deep learning, resulting in diverse fog nodes with different capabilities. Fog-based applications in healthcare can also leverage fog computing, which can analyze the information in real time and send an alert if necessary. Fog-aided IoT networks transfer CPU-intensive and delay-sensitive operations to a nearby fog node to overcome the limitations of real-time data transfer.
- Smart resource allocation and intelligent self-configuration are necessary for 6G wireless networks to optimize network resources and user satisfaction. Therefore, a collaborative resource management mechanism with multiple fog nodes and a cognitive context-aware resource allocation framework could meet the requirements for 6G wireless networks. Due to the enormous complexity of resource management scenarios, there is an increased requirement to implement ML algorithms in resource management in 6G wireless networks [12]. ML-based solutions allow direct access to data without using statistical models. The ML algorithms are also more robust to system parameters because they can utilize all information related to the data. Similarly, the

intractable problem, such as the NP-hard problem, could be tackled with affordable computational complexity using modern ML techniques.

- ML methods can address the complex computation requirements in 6G wireless networks. Mobile cloud computing (MCC) and mobile edge computing (MEC) can be used to transfer computationally intensive tasks from a mobile device to an adjacent edge or fog server to reduce latency and achieve computation efficiency [98]. A MEC approach allows mobile devices to delegate computationally-intensive tasks to a nearby server, which can be a simple node at the base station, a vehicle, or some other device. It involves real-time offloading decisions on a large scale. Traditional task offloading techniques are based on heuristic algorithms, game theory, and optimizations that demand a lot of processing and computational resources to execute the offloading decisions. Thus, ML algorithms, instead, offer several advantages over traditional approaches.
- Edge computing capabilities in 6G networks allow sensory data to be processed and analyzed near the source, significantly reducing the volume of transmitted data in 6G wireless networks. Hybrid ML-based solutions such as Fog-Learning integrated with edge computing could reduce the overall end-to-end system delay and the communication cost to address the task offloading and latency minimization in 6G applications. The centralized learning approach aggregates data from multiple devices into a central server. This enables centralized model training and global decision-making, which is well-suited for scenarios requiring massive data collection, extensive computations, and global optimization objectives, including network resource allocation, spectrum management, network security, and traffic prediction. Nevertheless, it raises concerns regarding data privacy and security.

  On the other hand, distributed learning excels in privacy-sensitive contexts, such as autonomous vehicles, the Internet of Medical Things (IoMT), personalized services, and extended reality applications in 6G networks. It emphasizes decentralized decision-making and fosters collaboration among nodes to create a global model. The DRL algorithms aggregate data to carry out centralized training. DRL models with federated training schemes may not significantly converge in large-scale MEC networks. However, designing the DRL model for a dynamic MEC network using frequent accessibility is still challenging for maintaining system stability.
- Massive self-organizing and self-healing devices will be vital aspects of 6G networks, which need extensive computing power. GPUs currently used in wireless networks do not meet next-generation wireless power efficiency requirements. Consequently, developing a scalable network architecture considering energy efficiency in 6G applications is necessary. Further, industry X.O applications and all the devices in IoTs, IoEV, IoMT, and V2X have sensors deployed everywhere, posing open research challenges regarding energy efficiency in wireless connectivity [119].

On the other hand, the extensive use of energy harvesting techniques and service requirements for 6G networks make it necessary for security configurations to be adaptable to available energy and network security risks. This can increase network security and network performance simultaneously [120]. Nevertheless, as network security threats are frequently anonymous, ML solutions combined with blockchain technology may represent an efficient technique for 6G networks to enhance security protection while utilizing the available power.

As a result of these lessons, we discuss future challenges in more detail and identify research directions in the next section.

V. OPEN ISSUES AND FUTURE TRENDS

This section comprises some of the open challenges and future trends that wireless networks can face while employing machine learning techniques in 6G communication systems:

A. CONTEXT-AWARE SMART RESOURCE ALLOCATION

There are congestion issues and high signaling overhead associated with the current random access (RA) allocation techniques, despite their support for mass machine-type communication (mMTC) [147]. A context-aware smart resource allocation is required for 6G wireless networks that support E2E connectivity. Furthermore, smart resource allocation is challenging for emerging 6G wireless applications such as tactile internet, extended reality, and the internet of vehicles due to strict latency requirements and seamless user experiences. Fast uplink grant (FUG) allocation introduced by 3GPP integrated with ML techniques could be an interesting research direction employed in modern Intelligent Transport systems (ITS) for latency-aware and reliable IoV applications under strict QoE constraints [148]. Moreover, machine learning models can accurately predict the traffic in V2X communications to avoid random allocation behavior. It can reduce signaling overhead and eliminate collisions, which allows IoV devices to consume minimal energy. Moreover, selecting appropriate exploration rates for V2X communication is also an attractive open research problem in 6G applications.

B. ML FOR GREEN COMMUNICATION

Research achievements in AI-based green communication services in the 6G era still need to be translated into practical applications [149], [150]. ML methods can be used in applications involving uncontrollable and predictable energy sources such as solar, wind, tide, and other renewable and partially controllable energy sources, including RF energy. An analysis of the relationship between uncontrollable but predictable energy harvesting technologies and the future harvesting power can also be conducted using ML methods [151], [152]. ML-assisted solutions are increasingly compelling for mapping complex relationships between current network traces and policies for future transmission as terrestrial network transmission policies change over time. These solutions could be extended to the Integrated Air-Ground-Underwater Network scenario [153].

C. PRIVACY-AWARE FLYING BASE STATIONS

In 6G networks, massive machine-type communication (mMTC) devices with ultra-dense and heterogeneous connectivity have attracted the research community to address the issues associated with the users' privacy and data security. As an aerial base station, a UAV serves as a space-air-terrestrial-sea communication channel to provide 6G services in areas without communication infrastructure and serves as mobile relays or data acquisition devices [154], [155]. ML techniques must be integrated with UAVs in 6G applications to support ultra-low latency communications for mMTC devices.

The UAV-based flying base station receives updates from the ML-trained models without altering or accessing any private information. Using D2D communication, UAVs can transmit the trained ML model to nearby devices for propagation to other clients. In UAV-assisted mMTC, recent strategies to train ML models, including transfer learning and federated learning, would be of interest as these methods allow the sharing of locally trained models while preventing the sharing of sensitive data to preserve the privacy of learners [156]. Further, it accommodates mMTC devices, ensures data privacy, and supports sustainable energy infrastructure design.

D. FOG LEARNING

Modern wireless applications in 6G networks require heterogeneous computation capabilities across devices, which poses several challenges to employing conventional federated learning and motivates device-to-device (D2D) intelligence in fog learning [157]. Fog learning (FogL) is an emerging paradigm that leverages the architecture of fog computing to execute machine learning tasks [158]. FogL entails extended federated learning to manage heterogeneous computation devices incorporating the fog network [159]. It assists intelligent model training through D2D communications at various network layers.

This hybrid learning technique intelligently distributes the ML training model across various nodes, including edge devices and cloud servers, to achieve optimized performance for computation capacity and local data distributions [160]. While considering the communication heterogeneity, multi-layer architecture of large-scale learning, privacy assumptions in D2D connectivity, and joint performance metrics as design parameters for developing Network-Aware Machine Learning tasks, federated learning has the limitations to be a suitable candidate for this solution. However, FogL-based solutions could be an interesting direction to address the task offloading and latency minimization in 6G applications.

E. MOBILE EDGE LEARNING

Edge computing capabilities in 6G networks allow sensory data to be processed and analyzed near the source, significantly reducing the transmitted data volume [161]. Some technical challenges are still involved in the practical implementation of mobile edge learning. In general, ML-training tasks require intensive computations, while edge devices are typically small and have limited computation and communication resources. The performance of mobile edge learning is inherently constrained by both connectivity and processing at edge devices. Further, mobile edge learning also encounters the "straggler's dilemma" in a heterogonous environment which limits the training time of ML models. Data samples at these edge devices may also not be independent and identically distributed (non-IID), compromising the effectiveness and efficiency of distributed training. However, a federated edge learning-based system could be a promising candidate to achieve computation efficiency in privacy and latency-aware 6G applications [162].

F. EXTENDED REALITY IN HEALTHCARE

Internet of Medical Things (IoMT) will become more valuable with the 6G communication network [163]. It will also have improved security features, allowing users to share sensitive information without any risk of interception or manipulation. Furthermore, extended reality (XR)-assisted teleoperation has demonstrated its ability to improve operational efficiency in healthcare-related complex scenarios [164]. A novel type of traffic with particular QoS requirements is introduced by the multisensory XR robots for which 6G network connectivity is anticipated in the coming years. ML-empowered solutions for cellular-connected VR and UAV networks in remote surgery using XR could be an interesting research direction for future applications.

G. INTERNET OF SENSES (IoS)

With the development of 6G, the desire for more advanced use cases involving the physical and digital worlds will be even more evident [165]. For instance, our senses will be expanded beyond our bodies through the Internet of Senses (IoS). Providing cost-effective and trustworthy solutions for these use cases will be possible through AI-enabled intelligent networks. The 6G networks will provide immersive communication in the IoS, enabling full telepresence and eliminating distance as a barrier to communication [166]. With the help of AI, personalized, immersive devices that can interact precisely with the human body will allow access to experiences and actions that are far away to improve human communications. Additionally, 6G networks will enable completely new modes of communication with strict controls over access and identity.

VI. CONCLUSION

This article presents a comprehensive survey of recent work on ML-enabled 6G wireless networks. The state-of-the-art

ML techniques, such as supervised and unsupervised machine learning, Deep Learning, Reinforcement Learning, Deep Reinforcement Learning, and Federated Learning for resource management applications in 6G networks, are reviewed. It discusses the implementation of ML algorithms in three network categories including device-to-device networks, vehicular networks, and Fog-Radio Access Networks. Furthermore, it summarizes the ML-based solutions to address the technical challenges in terms of resource allocation, task offloading, and handover management. Additionally, a comprehensive summary of ML-assisted methods to improve energy efficiency and reduce latency in 6G wireless networks is given. The aforementioned technical challenges and performance metrics will countenance the careful design of distributed ML architecture to circumvent the diverse resource optimization issues that will inevitably frost 6G networks.

This study concludes with the motivation and insight to leverage ML techniques in intelligent resource management for self-healing and self-configuration of 6G networks towards an end-to-end connected sustainable world. Finally, we highlight the open issues, challenges, and possible solutions with future research trends in the context of ML-enabled 6G wireless applications. These future research trends are anticipated to provide a new perspective to consider novel ML techniques in the design guidelines of emerging 6G wireless networks to automate network processes, analyze big data to make smart decisions and realize intelligent edge, fog, and cloud nodes with the ultimate goal of achieving seamless E2E connectivity.

REFERENCES

[1] A. Mazloomi, H. Sami, J. Bentahar, H. Otrok, and A. Mourad, "Reinforcement learning framework for server placement and workload allocation in multiaccess edge computing," *IEEE Internet Things J.*, vol. 10, no. 2, pp. 1376–1390, Jan. 2023, doi: 10.1109/JIOT.2022.3205051.

[2] H. R. Chi, C. K. Wu, N.-F. Huang, K.-F. Tsang, and A. Radwan, "A survey of network automation for industrial Internet-of-Things toward industry 5.0," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 2065–2077, Feb. 2023, doi: 10.1109/TII.2022.3215231.

[3] S. A. Khowaja, P. Khuwaja, K. Dev, I. H. Lee, W. U. Khan, W. Wang, N. M. F. Qureshi, and M. Magarini, "A secure data sharing scheme in community segmented vehicular social networks for 6G," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 890–899, Jan. 2023, doi: 10.1109/TII.2022.3188963.

[4] H. R. Chi, A. Radwan, N.-F. Huang, and K. F. Tsang, "Guest editorial: Next-generation network automation for industrial Internet-of-Things in industry 5.0," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 2062–2064, Feb. 2023, doi: 10.1109/TII.2022.3216903.

[5] Z. M. Fadlullah, B. Mao, and N. Kato, "Balancing QoS and security in the edge: Existing practices, challenges, and 6G opportunities with machine learning," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2419–2448, 4th Quart., 2022, doi: 10.1109/COMST.2022.3191697.

[6] L. Song, X. Hu, G. Zhang, P. Spachos, K. N. Plataniotis, and H. Wu, "Networking systems of AI: On the convergence of computing and communications," *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20352–20381, Oct. 2022, doi: 10.1109/JIOT.2022.3172270.

[7] M. R. Mahmood, M. A. Matin, P. Sarigiannidis, and S. K. Goudos, "A comprehensive review on artificial intelligence/machine learning algorithms for empowering the future IoT toward 6G era," *IEEE Access*, vol. 10, pp. 87535–87562, 2022, doi: 10.1109/ACCESS.2022.3199689.

[8] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019, doi: 10.1109/mcom.2019.1900271.

[9] Z. Yang, M. Chen, K.-K. Wong, H. V. Poor, and S. Cui, "Federated learning for 6G: Applications, challenges, and opportunities," *Engineering*, vol. 8, pp. 33–41, Jan. 2022, doi: 10.1016/j.eng.2021.12.002.

[10] J. Wang, J. Liu, J. Li, and N. Kato, "Artificial intelligence-assisted network slicing: Network assurance and service provisioning in 6G," *IEEE Veh. Technol. Mag.*, vol. 18, no. 1, pp. 49–58, Mar. 2023, doi: 10.1109/MVT.2022.3228399.

[11] W. Kim, Y. Ahn, J. Kim, and B. Shim, "Towards deep learning-aided wireless channel estimation and channel state information feedback for 6G," *J. Commun. Netw.*, vol. 25, no. 1, pp. 61–75, Feb. 2023, doi: 10.23919/JCN.2022.000037.

[12] E. Coronado, R. Behravesh, T. Subramanya, A. Fernàndez-Fernàndez, M. S. Siddiqui, X. Costa-Pérez, and R. Riggio, "Zero touch management: A survey of network automation solutions for 5G and 6G networks," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2535–2578, 4th Quart., 2022, doi: 10.1109/COMST.2022.3212586.

[13] A. Goldsmith, D. Gündüz, H. V. Poor, and Y. C. Eldar, "Machine learning for wireless networks," in *Machine Learning and Wireless Communications*, Cambridge, U.K.: Cambridge Univ. Press, 2022, pp. 21–22.

[14] C.-X. Wang, M. D. Renzo, S. Stanczak, S. Wang, and E. G. Larsson, "Artificial intelligence enabled wireless networking for 5G and beyond: Recent advances and future challenges," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 16–23, Feb. 2020, doi: 10.1109/mwc.001.1900292.

[15] M. Lin and Y. Zhao, "Artificial intelligence-empowered resource management for future wireless communications: A survey," *China Commun.*, vol. 17, no. 3, pp. 58–77, Mar. 2020, doi: 10.23919/JCC.2020.03.006.

[16] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-intelligence-enabled intelligent 6G networks," *IEEE Netw.*, vol. 34, no. 6, pp. 272–280, Nov. 2020, doi: 10.1109/mnet.011.2000195.

[17] A. Salh, L. Audah, N. S. M. Shah, A. Alhammadi, Q. Abdullah, Y. H. Kim, S. A. Al-Gailani, S. A. Hamzah, B. A. F. Esmail, and A. A. Almohammedi, "A survey on deep learning for ultra-reliable and low-latency communications challenges on 6G wireless systems," *IEEE Access*, vol. 9, pp. 55098–55131, 2021, doi: 10.1109/access.2021.3069707.

[18] L. Lei, Y. Tan, K. Zheng, S. Liu, K. Zhang, and X. Shen, "Deep reinforcement learning for autonomous Internet of Things: Model, applications and challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1722–1760, 3rd Quart., 2020, doi: 10.1109/comst.2020.2988367.

[19] B. Brik, K. Boutiba, and A. Ksentini, "Deep learning for B5G open radio access network: Evolution, survey, case studies, and challenges," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 228–250, 2022, doi: 10.1109/ojcoms.2022.3146618.

[20] M. S. Mollel, A. I. Abubakar, M. Ozturk, S. F. Kaijage, M. Kisangiri, S. Hussain, M. A. Imran, and Q. H. Abbasi, "A survey of machine learning applications to handover management in 5G and beyond," *IEEE Access*, vol. 9, pp. 45770–45802, 2021, doi: 10.1109/access.2021.3067503.

[21] U. M. Malik, M. A. Javed, S. Zeadally, and S. u. Islam, "Energy-efficient fog computing for 6G-enabled massive IoT: Recent trends and future opportunities," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14572–14594, Aug. 2022, doi: 10.1109/jiot.2021.3068056.

[22] A. Munir, J. Kwon, J. H. Lee, J. Kong, E. Blasch, A. J. Aved, and K. Muhammad, "FogSurv: A fog-assisted architecture for urban surveillance using artificial intelligence and data fusion," *IEEE Access*, vol. 9, pp. 111938–111959, 2021, doi: 10.1109/access.2021.3102598.

[23] M. S. M. Gismalla, A. I. Azmi, M. R. B. Salim, M. F. L. Abdullah, F. Iqbal, W. A. Mabrouk, M. B. Othman, A. Y. I. Ashyap, and A. S. M. Supa'at, "Survey on device to device (D2D) communication for 5GB/6G networks: Concept, applications, challenges, and future directions," *IEEE Access*, vol. 10, pp. 30792–30821, 2022, doi: 10.1109/access.2022.3160215.

[24] E. Gures, I. Shayea, M. Ergen, M. H. Azmi, and A. A. El-Saleh, "Machine learning-based load balancing algorithms in future heterogeneous networks: A survey," *IEEE Access*, vol. 10, pp. 37689–37717, 2022, doi: 10.1109/access.2022.3161511.

[25] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6G: Vision, enabling technologies, and applications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 5–36, Jan. 2022, doi: 10.1109/jsac.2021.3126076.

[26] J. Kaur, M. A. Khan, M. Iftikhar, M. Imran, and Q. E. U. Haq, "Machine learning techniques for 5G and beyond," *IEEE Access*, vol. 9, pp. 23472–23488, 2021, doi: 10.1109/access.2021.3051557.

[27] H. Dahrouj, R. Alghamdi, H. Alwazani, S. Bahanshal, A. A. Ahmad, A. Faisal, R. Shalabi, R. Alhadrami, A. Subasi, M. T. Al-Nory, O. Kittaneh, and J. S. Shamma, "An overview of machine learning-based techniques for solving optimization problems in communications and signal processing," *IEEE Access*, vol. 9, pp. 74908–74938, 2021, doi: 10.1109/access.2021.3079639.

[28] H. Fourati, R. Maaloul, and L. Chaari, "A survey of 5G network systems: Challenges and machine learning approaches," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 2, pp. 385–431, Feb. 2021, doi: 10.1007/s13042-020-01178-4.

[29] O. Nassef, W. Sun, H. Purmehdi, M. Tatipamula, and T. Mahmoodi, "A survey: Distributed machine learning for 5G and beyond," *Comput. Netw.*, vol. 207, Apr. 2022, Art. no. 108820, doi: 10.1016/j.comnet.2022.108820.

[30] I. A. Bartsiokas, P. K. Gkonis, D. I. Kaklamani, and I. S. Venieris, "ML-based radio resource management in 5G and beyond networks: A survey," *IEEE Access*, vol. 10, pp. 83507–83528, 2022, doi: 10.1109/access.2022.3196657.

[31] M. A. Ridwan, N. A. M. Radzi, F. Abdullah, and Y. E. Jalil, "Applications of machine learning in networking: A survey of current issues and future challenges," *IEEE Access*, vol. 9, pp. 52523–52556, 2021, doi: 10.1109/access.2021.3069210.

[32] H. Tran-Dang, S. Bhardwaj, T. Rahim, A. Musaddiq, and D.-S. Kim, "Reinforcement learning based resource management for fog computing environment: Literature review, challenges, and open issues," *J. Commun. Netw.*, vol. 24, no. 1, pp. 83–98, Feb. 2022, doi: 10.23919/jcn.2021.000041.

[33] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine learning for resource management in cellular and IoT networks: Potentials, current solutions, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1251–1275, 2nd Quart., 2020, doi: 10.1109/comst.2020.2964534.

[34] M. Noor-A-Rahim, Z. Liu, H. Lee, G. G. M. N. Ali, D. Pesch, and P. Xiao, "A survey on resource allocation in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 701–721, Feb. 2022, doi: 10.1109/tits.2020.3019322.

[35] T. T. Thanh Le and S. Moh, "Comprehensive survey of radio resource allocation schemes for 5G V2X communications," *IEEE Access*, vol. 9, pp. 123117–123133, 2021, doi: 10.1109/access.2021.3109894.

[36] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine learning for 6G wireless networks: Carrying forward enhanced bandwidth, massive access, and ultrareliable/low-latency service," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 122–134, Dec. 2020, doi: 10.1109/mvt.2020.3019650.

[37] B. Mao, F. Tang, Y. Kawamoto, and N. Kato, "AI models for green communications towards 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 210–247, 1st Quart., 2022, doi: 10.1109/comst.2021.3130901.

[38] M. A. Hossain, A. R. Hossain, and N. Ansari, "AI in 6G: Energy-efficient distributed machine learning for multilayer heterogeneous networks," *IEEE Netw.*, vol. 36, no. 6, pp. 84–91, Nov. 2022, doi: 10.1109/MNET.104.2100422.

[39] T. B. Ahammed, R. Patgiri, and S. Nayak, "A vision on the artificial intelligence for 6G communication," *ICT Exp.*, vol. 9, no. 2, pp. 197–210, Apr. 2023, doi: 10.1016/j.icte.2022.05.005.

[40] P. S. Bouzinis, P. D. Diamantoulakis, and G. K. Karagiannidis, "Wireless federated learning (WFL) for 6G networks4Part I: Research challenges and future trends," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 3–7, Jan. 2022, doi: 10.1109/lcomm.2021.3121071.

[41] B. Ji, Y. Wang, K. Song, C. Li, H. Wen, V. G. Menon, and S. Mumtaz, "A survey of computational intelligence for 6G: Key technologies, applications and trends," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 7145–7154, Oct. 2021, doi: 10.1109/tii.2021.3052531.

[42] V. P. Rekkas, S. Sotiroudis, P. Sarigiannidis, S. Wan, G. K. Karagiannidis, and S. K. Goudos, "Machine learning in beyond 5G/6G networks—State-of-the-art and future trends," *Electronics*, vol. 10, no. 22, p. 2786, Nov. 2021, doi: 10.3390/electronics10222786.

[43] I. I. Ioannou, C. Christophorou, V. Vassiliou, M. Lestas, and A. Pitsillides, "Dynamic D2D communication in 5G/6G using a distributed AI framework," *IEEE Access*, vol. 10, pp. 62772–62799, 2022, doi: 10.1109/ACCESS.2022.3182388.

[44] S. Aslam, F. Alam, S. F. Hasan, and M. A. Rashid, "A machine learning approach to enhance the performance of D2D-enabled clustered networks," *IEEE Access*, vol. 9, pp. 16114–16132, 2021, doi: 10.1109/access.2021.3053045.

[45] I. Ioannou, V. Vassiliou, C. Christophorou, and A. Pitsillides, "Distributed artificial intelligence solution for D2D communication in 5G networks," *IEEE Syst. J.*, vol. 14, no. 3, pp. 4232–4241, Sep. 2020, doi: 10.1109/JSYST.2020.2979044.

[46] W. Wu, R. Liu, Q. Yang, and T. Q. S. Quek, "Learning-based robust resource allocation for D2D underlaying cellular network," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6731–6745, Aug. 2022, doi: 10.1109/TWC.2022.3152260.

[47] R. Bajpai, S. Chakraborty, and N. Gupta, "Adapting deep learning for content caching frameworks in device-to-device environments," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 912–920, 2022, doi: 10.1109/ojcoms.2022.3175927.

[48] N. Simmons, S. B. F. Gomes, M. D. Yacoub, O. Simeone, S. L. Cotton, and D. E. Simmons, "AI-based channel prediction in D2D links: An empirical validation," *IEEE Access*, vol. 10, pp. 65459–65472, 2022, doi: 10.1109/access.2022.3182713.

[49] M. Sengly, K. Lee, and J.-R. Lee, "Joint optimization of spectral efficiency and energy harvesting in D2D networks using deep neural network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8361–8366, Aug. 2021, doi: 10.1109/tvt.2021.3055205.

[50] W. Lee and R. Schober, "Deep learning-based resource allocation for device-to-device communication," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5235–5250, Jul. 2022, doi: 10.1109/TWC.2021.3138733.

[51] L. Feng, Z. Yang, Y. Yang, X. Que, and K. Zhang, "Smart mode selection using online reinforcement learning for VR broadband broadcasting in D2D assisted 5G HetNets," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 600–611, Jun. 2020, doi: 10.1109/tbc.2020.2977577.

[52] M. S. Al-Abiad, M. Z. Hassan, and M. J. Hossain, "A joint reinforcement-learning enabled caching and cross-layer network code in F-RAN with D2D communications," *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4400–4416, Jul. 2022, doi: 10.1109/TCOMM.2022.3168058.

[53] D. Ron and J.-R. Lee, "DRL-based sum-rate maximization in D2D communication underlaid uplink cellular networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11121–11126, Oct. 2021, doi: 10.1109/tvt.2021.3106398.

[54] M. V. da Silva, S. Montejo-Sánchez, R. D. Souza, H. Alves, and T. Abrão, "D2D assisted Q-learning random access for NOMA-based MTC networks," *IEEE Access*, vol. 10, pp. 30694–30706, 2022, doi: 10.1109/access.2022.3160156.

[55] R. Cheng, Y. Sun, Y. Liu, L. Xia, D. Feng, and M. A. Imran, "Blockchain-empowered federated learning approach for an intelligent and reliable D2D caching scheme," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 7879–7890, Jun. 2022, doi: 10.1109/jiot.2021.3103107.

[56] Q. Guo, F. Tang, and N. Kato, "Federated reinforcement learning-based resource allocation for D2D-aided digital twin edge networks in 6G industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 19, no. 5, pp. 7228–7236, May 2023, doi: 10.1109/tii.2022.3227655.

[57] G. Raja, M. Begum, S. Gurumoorthy, D. S. Rajendran, P. Srividya, K. Dev, and N. M. F. Qureshi, "AI-empowered trajectory anomaly detection and classification in 6G-V2X," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4599–4607, Apr. 2023, doi: 10.1109/TITS.2022.3197446.

[58] J. He, K. Yang, and H.-H. Chen, "6G cellular networks and connected autonomous vehicles," *IEEE Netw.*, vol. 35, no. 4, pp. 255–261, Jul. 2021, doi: 10.1109/MNET.011.2000541.

[59] M. Ergen, F. Inan, O. Ergen, I. Shayea, M. F. Tuysuz, A. Azizan, N. K. Ure, and M. Nekovee, "Edge on wheels with OMNIBUS networking for 6G technology," *IEEE Access*, vol. 8, pp. 215928–215942, 2020, doi: 10.1109/ACCESS.2020.3038233.

[60] F. Falahatraftar, S. Pierre, and S. Chamberland, "A centralized and dynamic network congestion classification approach for heterogeneous vehicular networks," *IEEE Access*, vol. 9, pp. 122284–122298, 2021, doi: 10.1109/access.2021.3108425.

[61] A. H. Sodhro, G. H. Sodhro, M. Guizani, S. Pirbhulal, and A. Boukerche, "AI-enabled reliable channel modeling architecture for fog computing vehicular networks," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 14–21, Apr. 2020, doi: 10.1109/mwc.001.1900311.

[62] Z. Liu, P. Dai, H. Xing, Z. Yu, and W. Zhang, "A distributed algorithm for task offloading in vehicular networks with hybrid fog/cloud computing," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 7, pp. 4388–4401, Jul. 2022, doi: 10.1109/TSMC.2021.3097005.

[63] S. Shen, C. Yu, K. Zhang, and S. Ci, "Adaptive artificial intelligence for resource-constrained connected vehicles in cybertwin-driven 6G network," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16269–16278, Nov. 2021, doi: 10.1109/JIOT.2021.3101231.

[64] S. K. Tayyaba, H. A. Khattak, A. Almogren, M. A. Shah, I. U. Din, I. Alkhalifa, and M. Guizani, "5G vehicular network resource management for improving radio access through machine learning," *IEEE Access*, vol. 8, pp. 6792–6800, 2020, doi: 10.1109/ACCESS.2020.2964697.

[65] M. Ibrar, A. Akbar, S. R. U. Jan, M. A. Jan, L. Wang, H. Song, and N. Shah, "ARTNet: AI-based resource allocation and task offloading in a reconfigurable Internet of Vehicular networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 1, pp. 67–77, Jan. 2022, doi: 10.1109/TNSE.2020.3047454.

[66] H. Xiang, S. Yan, and M. Peng, "A realization of fog-RAN slicing via deep reinforcement learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2515–2527, Apr. 2020, doi: 10.1109/TWC.2020.2965927.

[67] X. Wang, Y. Wang, Q. Cui, K.-C. Chen, and W. Ni, "Machine learning enables radio resource allocation in the downlink of ultra-low latency vehicular networks," *IEEE Access*, vol. 10, pp. 44710–44723, 2022, doi: 10.1109/ACCESS.2022.3168986.

[68] J. Zhao, M. Kong, Q. Li, and X. Sun, "Contract-based computing resource management via deep reinforcement learning in vehicular fog computing," *IEEE Access*, vol. 8, pp. 3319–3329, 2020, doi: 10.1109/ACCESS.2019.2963051.

[69] Y. Dai, D. Xu, K. Zhang, S. Maharjan, and Y. Zhang, "Deep reinforcement learning and permissioned blockchain for content caching in vehicular edge computing and networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4312–4324, Apr. 2020, doi: 10.1109/tvt.2020.2973705.

[70] G. Sun, G. O. Boateng, D. Ayepah-Mensah, G. Liu, and J. Wei, "Autonomous resource slicing for virtualized vehicular networks with D2D communications based on deep reinforcement learning," *IEEE Syst. J.*, vol. 14, no. 4, pp. 4694–4705, Dec. 2020, doi: 10.1109/jsyst.2020.2982857.

[71] X. Zhang, M. Peng, S. Yan, and Y. Sun, "Deep-reinforcement-learning-based mode selection and resource allocation for cellular V2X communications," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6380–6391, Jul. 2020, doi: 10.1109/jiot.2019.2962715.

[72] X. Li, L. Lu, W. Ni, A. Jamalipour, D. Zhang, and H. Du, "Federated multi-agent deep reinforcement learning for resource allocation of vehicle-to-vehicle communications," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8810–8824, Aug. 2022, doi: 10.1109/tvt.2022.3173057.

[73] Z. Zhou, H. Liao, B. Gu, S. Mumtaz, and J. Rodriguez, "Resource sharing and task offloading in IoT fog computing: A contract-learning approach," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 4, no. 3, pp. 227–240, Jun. 2020, doi: 10.1109/tetci.2019.2902869.

[74] Y. Jiang, H. Feng, F.-C. Zheng, D. Niyato, and X. You, "Deep learning-based edge caching in fog radio access networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8442–8454, Dec. 2020, doi: 10.1109/twc.2020.3022907.

[75] W. Wu, M. Peng, W. Chen, and S. Yan, "Unsupervised deep transfer learning for fault diagnosis in fog radio access networks," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8956–8966, Sep. 2020, doi: 10.1109/jiot.2020.2997187.

[76] W. Xia, X. Zhang, G. Zheng, J. Zhang, S. Jin, and H. Zhu, "The interplay between artificial intelligence and fog radio access networks," *China Commun.*, vol. 17, no. 8, pp. 1–13, Aug. 2020, doi: 10.23919/JCC.2020.08.001.

[77] Q. Fan, J. Bai, H. Zhang, Y. Yi, and L. Liu, "Delay-aware resource allocation in fog-assisted IoT networks through reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 7, pp. 5189–5199, Apr. 2022, doi: 10.1109/jiot.2021.3111079.

[78] S. Yan, M. Jiao, Y. Zhou, M. Peng, and M. Daneshmand, "Machine-learning approach for user association and content placement in fog radio access networks," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9413–9425, Oct. 2020, doi: 10.1109/jiot.2020.2973339.

[79] H. Xiang, M. Peng, Y. Sun, and S. Yan, "Mode selection and resource allocation in sliced fog radio access networks: A reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4271–4284, Apr. 2020, doi: 10.1109/tvt.2020.2972999.

[80] C. Fang, H. Xu, Y. Yang, Z. Hu, S. Tu, K. Ota, Z. Yang, M. Dong, Z. Han, F. R. Yu, and Y. Liu, "Deep-reinforcement-learning-based resource allocation for content distribution in fog radio access networks," *IEEE Internet Things J.*, vol. 9, no. 18, pp. 16874–16883, Sep. 2022, doi: 10.1109/jiot.2022.3146239.

[81] G. M. S. Rahman, M. Peng, S. Yan, and T. Dang, "Learning based joint cache and power allocation in fog radio access networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4401–4411, Apr. 2020, doi: 10.1109/tvt.2020.2975849.

[82] Y. Sun, M. Peng, Y. Ren, L. Chen, L. Yu, and S. Suo, "Harmonizing artificial intelligence with radio access networks: Advances, case study, and open issues," *IEEE Netw.*, vol. 35, no. 4, pp. 144–151, Jul. 2021, doi: 10.1109/mnet.011.2000656.

[83] L. Zhang, Y. Jiang, F.-C. Zheng, M. Bennis, and X. You, "Computation offloading and resource allocation in F-RANs: A federated deep reinforcement learning approach," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 97–102.

[84] Z. Yu, J. Hu, G. Min, Z. Wang, W. Miao, and S. Li, "Privacy-preserving federated deep learning for cooperative hierarchical caching in fog computing," *IEEE Internet Things J.*, vol. 9, no. 22, pp. 22246–22255, Nov. 2022, doi: 10.1109/jiot.2021.3081480.

[85] K. Mishra, G. N. V. Rajareddy, U. Ghugar, G. S. Chhabra, and A. H. Gandomi, "A collaborative computation and offloading for compute-intensive and latency-sensitive dependency-aware tasks in dew-enabled vehicular fog computing: A federated deep Q-learning approach," *IEEE Trans. Netw. Service Manage.*, early access, Jun. 5, 2023, doi: 10.1109/tnsm.2023.3282795.

[86] A. M. Elbir and S. Coleri, "Federated learning for channel estimation in conventional and RIS-assisted massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4255–4268, Jun. 2022, doi: 10.1109/twc.2021.3128392.

[87] M. Merluzzi, P. D. Lorenzo, and S. Barbarossa, "Wireless edge machine learning: Resource allocation and trade-offs," *IEEE Access*, vol. 9, pp. 45377–45398, 2021, doi: 10.1109/access.2021.3066559.

[88] L. Zhou, Y. Hong, S. Wang, R. Han, D. Li, R. Wang, and Q. Hao, "Learning centric wireless resource allocation for edge computing: Algorithm and experiment," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 1035–1040, Jan. 2021, doi: 10.1109/tvt.2020.3047149.

[89] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "LORM: Learning to optimize for resource management in wireless networks with few training samples," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 665–679, Jan. 2020, doi: 10.1109/twc.2019.2947591.

[90] M. A. Jamshed, F. Héliot, and T. W. C. Brown, "Unsupervised learning based emission-aware uplink resource allocation scheme for non-orthogonal multiple access systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7681–7691, Aug. 2021, doi: 10.1109/tvt.2021.3089898.

[91] R. Alkurd, I. Y. Abualhaol, and H. Yanikomeroglu, "Personalized resource allocation in wireless networks: An AI-enabled and big data-driven multi-objective optimization," *IEEE Access*, vol. 8, pp. 144592–144609, 2020, doi: 10.1109/access.2020.3014301.

[92] E. Kim, H.-H. Choi, H. Kim, J. Na, and H. Lee, "Optimal resource allocation considering non-uniform spatial traffic distribution in ultra-dense networks: A multi-agent reinforcement learning approach," *IEEE Access*, vol. 10, pp. 20455–20464, 2022, doi: 10.1109/access.2022.3152162.

[93] W. Ahsan, W. Yi, Z. Qin, Y. Liu, and A. Nallanathan, "Resource allocation in uplink NOMA-IoT networks: A reinforcement-learning approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5083–5098, Aug. 2021, doi: 10.1109/twc.2021.3065523.

[94] F. Tang, Y. Zhou, and N. Kato, "Deep reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5G HetNet," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2773–2782, Dec. 2020, doi: 10.1109/jsac.2020.3005495.

[95] M. S. Allahham, A. A. Abdellatif, N. Mhaisen, A. Mohamed, A. Erbad, and M. Guizani, "Multi-agent reinforcement learning for network selection and resource allocation in heterogeneous multi-RAT networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 2, pp. 1287–1300, Jun. 2022, doi: 10.1109/tccn.2022.3155727.

[96] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, and X. Costa-Pérez, "A machine learning approach to 5G infrastructure market optimization," *IEEE Trans. Mobile Comput.*, vol. 19, no. 3, pp. 498–512, Mar. 2020, doi: 10.1109/tmc.2019.2896950.

[97] Z. Shi, J. Liu, S. Zhang, and N. Kato, "Multi-agent deep reinforcement learning for massive access in 5G and beyond ultra-dense NOMA system," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3057–3070, May 2022, doi: 10.1109/twc.2021.3117859.

[98] Q. Zhang, H. Wen, Y. Liu, S. Chang, and Z. Han, "Federated-reinforcement-learning-enabled joint communication, sensing, and computing resources allocation in connected automated vehicles networks," *IEEE Internet Things J.*, vol. 9, no. 22, pp. 23224–23240, Nov. 2022, doi: 10.1109/jiot.2022.3188434.

[99] R. Zhong, X. Liu, Y. Liu, Y. Chen, and Z. Han, "Mobile reconfigurable intelligent surfaces for NOMA networks: Federated learning approaches," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 10020–10034, Nov. 2022, doi: 10.1109/twc.2022.3181747.

[100] P. Zeng, A. Liu, C. Zhu, T. Wang, and S. Zhang, "Trust-based multi-agent imitation learning for green edge computing in smart cities," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 3, pp. 1635–1648, Sep. 2022, doi: 10.1109/TGCN.2022.3172367.

[101] H. Wu, Z. Zhang, C. Guan, K. Wolter, and M. Xu, "Collaborate edge and cloud computing with distributed deep learning for smart city Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8099–8110, Sep. 2020, doi: 10.1109/jiot.2020.2996784.

[102] H. Liao, Y. Mu, Z. Zhou, M. Sun, Z. Wang, and C. Pan, "Blockchain and learning-based secure and intelligent task offloading for vehicular fog computing," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4051–4063, Jul. 2021, doi: 10.1109/tits.2020.3007770.

[103] I. Khan, X. Tao, G. M. S. Rahman, W. U. Rehman, and T. Salam, "Advanced energy-efficient computation offloading using deep reinforcement learning in MTC edge computing," *IEEE Access*, vol. 8, pp. 82867–82875, 2020, doi: 10.1109/access.2020.2991057.

[104] H. Zhou, T. Wu, H. Zhang, and J. Wu, "Incentive-driven deep reinforcement learning for content caching and D2D offloading," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2445–2460, Aug. 2021, doi: 10.1109/jsac.2021.3087232.

[105] F. Khoramnejad and M. Erol-Kantarci, "On joint offloading and resource allocation: A double deep Q-network approach," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 4, pp. 1126–1141, Dec. 2021, doi: 10.1109/tccn.2021.3116251.

[106] M. Nduwayezu, Q.-V. Pham, and W.-J. Hwang, "Online computation offloading in NOMA-based multi-access edge computing: A deep reinforcement learning approach," *IEEE Access*, vol. 8, pp. 99098–99109, 2020, doi: 10.1109/access.2020.2997925.

[107] Z. Cao, P. Zhou, R. Li, S. Huang, and D. Wu, "Multiagent deep reinforcement learning for joint multichannel access and task offloading of mobile-edge computing in industry 4.0," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6201–6213, Jul. 2020, doi: 10.1109/jiot.2020.2968951.

[108] L. Zang, X. Zhang, and B. Guo, "Federated deep reinforcement learning for online task offloading and resource allocation in WPC-MEC networks," *IEEE Access*, vol. 10, pp. 9856–9867, 2022, doi: 10.1109/access.2022.3144415.

[109] S. Yu, X. Chen, Z. Zhou, X. Gong, and D. Wu, "When deep reinforcement learning meets federated learning: Intelligent multitimescale resource management for multiaccess edge computing in 5G ultradense network," *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2238–2251, Feb. 2021, doi: 10.1109/jiot.2020.3026589.

[110] C. Pan, Z. Wang, H. Liao, Z. Zhou, X. Wang, M. Tariq, and S. Al-Otaibi, "Asynchronous federated deep reinforcement learning-based URLLC-aware computation offloading in space-assisted vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 7, pp. 7377–7389, Jul. 2022, doi: 10.1109/TITS.2022.3150756.

[111] D. Prerna, R. Tekchandani, N. Kumar, and S. Tanwar, "An energy-efficient cache localization technique for D2D communication in IoT environment," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4816–4829, Mar. 2021, doi: 10.1109/jiot.2020.3029168.

[112] B. Özbek, M. Pischella, and D. Le Ruyet, "Energy efficient resource allocation for underlaying multi-D2D enabled multiple-antennas communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6189–6199, Jun. 2020, doi: 10.1109/tvt.2020.2981409.

[113] J. Tang, H. Tang, X. Zhang, K. Cumanan, G. Chen, K.-K. Wong, and J. A. Chambers, "Energy minimization in D2D-assisted cache-enabled Internet of Things: A deep reinforcement learning approach," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5412–5423, Aug. 2020, doi: 10.1109/tii.2019.2954127.

[114] Z. Ji, A. K. Kiani, Z. Qin, and R. Ahmad, "Power optimization in device-to-device communications: A deep reinforcement learning approach with dynamic reward," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 508–511, Mar. 2021, doi: 10.1109/lwc.2020.3035898.

[115] T. Zhang, K. Zhu, and J. Wang, "Energy-efficient mode selection and resource allocation for D2D-enabled heterogeneous networks: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1175–1187, Feb. 2021, doi: 10.1109/twc.2020.3031436.

[116] X. Chen and G. Liu, "Energy-efficient task offloading and resource allocation via deep reinforcement learning for augmented reality in mobile edge networks," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10843–10856, Jul. 2021, doi: 10.1109/jiot.2021.3050804.

[117] Y. Liu, L. Jiang, Q. Qi, and S. Xie, "Energy-efficient space–air–ground integrated edge computing for Internet of Remote things: A federated DRL approach," *IEEE Internet Things J.*, vol. 10, no. 6, pp. 4845–4856, Mar. 2023, doi: 10.1109/jiot.2022.3220677.

[118] S. Zarandi and H. Tabassum, "Federated double deep Q-learning for joint delay and energy minimization in IoT networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.

[119] Y. Nie, J. Zhao, F. Gao, and F. R. Yu, "Semi-distributed resource management in UAV-aided MEC systems: A multi-agent federated reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 13162–13173, Dec. 2021, doi: 10.1109/tvt.2021.3118446.

[120] S. Wang, M. Chen, C. Yin, W. Saad, C. S. Hong, S. Cui, and H. V. Poor, "Federated learning for task and resource allocation in wireless high-altitude balloon networks," *IEEE Internet Things J.*, vol. 8, no. 24, pp. 17460–17475, Dec. 2021, doi: 10.1109/jiot.2021.3080078.

[121] S. A. Khowaja, K. Dev, P. Khowaja, and P. Bellavista, "Toward energy-efficient distributed federated learning for 6G networks," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 34–40, Dec. 2021, doi: 10.1109/mwc.012.2100153.

[122] R. Luo, W. Ni, H. Tian, and J. Cheng, "Federated deep reinforcement learning for RIS-assisted indoor multi-robot communication systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 11, pp. 12321–12326, Nov. 2022, doi: 10.1109/tvt.2022.3190557.

[123] M. A. Jamshed, K. Ali, Q. H. Abbasi, M. A. Imran, and M. Ur-Rehman, "Challenges, applications, and future of wireless sensors in Internet of Things: A review," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5482–5494, Mar. 2022, doi: 10.1109/JSEN.2022.3148128.

[124] J. Wang, K. Zhu, and E. Hossain, "Green Internet of Vehicles (IoV) in the 6G era: Toward sustainable vehicular communications and networking," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 391–423, Mar. 2022, doi: 10.1109/TGCN.2021.3127923.

[125] B. Mao, Y. Kawamoto, and N. Kato, "AI-based joint optimization of QoS and security for 6G energy harvesting Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7032–7042, Aug. 2020, doi: 10.1109/jiot.2020.2982417.

[126] E. Eldeeb, M. Shehab, and H. Alves, "A learning-based fast uplink grant for massive IoT via support vector machines and long short-term memory," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3889–3898, Mar. 2022, doi: 10.1109/jiot.2021.3101978.

[127] X. Cai, X. Mo, J. Chen, and J. Xu, "D2D-enabled data sharing for distributed machine learning at wireless network edge," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1457–1461, Sep. 2020, doi: 10.1109/lwc.2020.2993837.

[128] S. Wang, T. Lv, X. Zhang, Z. Lin, and P. Huang, "Learning-based multi-channel access in 5G and beyond networks with fast time-varying channels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5203–5218, May 2020, doi: 10.1109/tvt.2020.2980861.

[129] X. Wang, R. Li, C. Wang, X. Li, T. Taleb, and V. C. M. Leung, "Attention-weighted federated deep reinforcement learning for device-to-device assisted heterogeneous collaborative edge caching," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 154–169, Jan. 2021, doi: 10.1109/jsac.2020.3036946.

[130] H. Alsulami, S. H. Serbaya, E. H. Abualsauod, A. M. Othman, A. Rizwan, and A. Jalali, "A federated deep learning empowered resource management method to optimize 5G and 6G quality of services (QoS)," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–9, Mar. 2022, doi: 10.1155/2022/1352985.

[131] X. Wang, C. Wang, X. Li, V. C. M. Leung, and T. Taleb, "Federated deep reinforcement learning for Internet of Things with decentralized cooperative edge caching," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9441–9455, Oct. 2020, doi: 10.1109/jiot.2020.2986803.

[132] S. Messaoud, A. Bradai, O. B. Ahmed, P. T. A. Quang, M. Atri, and M. S. Hossain, "Deep federated Q-learning-based network slicing for industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5572–5582, Aug. 2021, doi: 10.1109/tii.2020.3032165.

[133] Y. Liu, H. Zhou, Y. Deng, and A. Nallanathan, "Channel access optimization in unlicensed spectrum for downlink URLLC: Centralized and federated DRL approaches," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 7, pp. 2208–2222, Jul. 2023, doi: 10.1109/jsac.2023.3280982.

[134] A. Tian, B. Feng, H. Zhou, Y. Huang, K. Sood, S. Yu, and H. Zhang, "Efficient federated DRL-based cooperative caching for mobile edge networks," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 1, pp. 246–260, Mar. 2023, doi: 10.1109/tnsm.2022.3198074.

[135] Z. Ali, L. Giupponi, M. Miozzo, and P. Dini, "Multi-task learning for efficient management of beyond 5G radio access network architectures," *IEEE Access*, vol. 9, pp. 158892–158907, 2021, doi: 10.1109/access.2021.3130740.

[136] M.-T. Nguyen and S. Kwon, "Machine learning–based mobility robustness optimization under dynamic cellular networks," *IEEE Access*, vol. 9, pp. 77830–77844, 2021, doi: 10.1109/access.2021.3083554.

[137] S. Khosravi, H. Shokri-Ghadikolaei, and M. Petrova, "Learning-based handover in mobile millimeter-wave networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 2, pp. 663–674, Jun. 2021, doi: 10.1109/TCCN.2020.3030964.

[138] Y. Sun, W. Jiang, G. Feng, P. V. Klaine, L. Zhang, M. A. Imran, and Y.-C. Liang, "Efficient handover mechanism for radio access network slicing by exploiting distributed learning," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 4, pp. 2620–2633, Dec. 2020, doi: 10.1109/tnsm.2020.3031079.

[139] M. Sana, A. De Domenico, W. Yu, Y. Lostanlen, and E. Calvanese Strinati, "Multi-agent reinforcement learning for adaptive user association in dynamic mmWave networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6520–6534, Oct. 2020, doi: 10.1109/twc.2020.3003719.

[140] Y. Koda, K. Nakashima, K. Yamamoto, T. Nishio, and M. Morikura, "Handover management for mmWave networks with proactive performance prediction using camera images and deep reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 2, pp. 802–816, Jun. 2020, doi: 10.1109/tccn.2019.2961655.

[141] M. S. Mollel, S. Kaijage, and M. Kisangiri, "Deep reinforcement learning based handover management for millimeter wave communication," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 2, 2021, Art. no. 0120298.

[142] K. Tan, D. Bremner, J. Le Kernec, Y. Sambo, L. Zhang, and M. A. Imran, "Intelligent handover algorithm for vehicle-to-network communications with double-deep Q-learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7848–7862, Jul. 2022, doi: 10.1109/tvt.2022.3169804.

[143] D. Guo, L. Tang, X. Zhang, and Y.-C. Liang, "Joint optimization of handover control and power allocation based on multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13124–13138, Nov. 2020, doi: 10.1109/tvt.2020.3020400.

[144] Y. Lin, Z. Zhang, Y. Huang, J. Li, F. Shu, and L. Hanzo, "Heterogeneous user-centric cluster migration improves the connectivity-handover trade-off in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16027–16043, Dec. 2020, doi: 10.1109/tvt.2020.3041521.

[145] Y. Cao, S.-Y. Lien, Y.-C. Liang, K.-C. Chen, and X. Shen, "User access control in open radio access networks: A federated deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3721–3736, Jun. 2022, doi: 10.1109/TWC.2021.3123500.

[146] S. Hosseinalipour, S. S. Azam, C. G. Brinton, N. Michelusi, V. Aggarwal, D. J. Love, and H. Dai, "Multi-stage hybrid federated learning over large-scale D2D-enabled fog networks," *IEEE/ACM Trans. Netw.*, vol. 30, no. 4, pp. 1569–1584, Aug. 2022, doi: 10.1109/tnet.2022.3143495.

[147] H. F. Alhashimi, M. N. Hindia, K. Dimyati, E. B. Hanafi, N. Safie, F. Qamar, K. Azrin, and Q. N. Nguyen, "A survey on resource management for 6G heterogeneous networks: Current research, future trends, and challenges," *Electronics*, vol. 12, no. 3, p. 647, Jan. 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/3/647

[148] L. Liu, Z. Lu, L. Wang, Y. Chen, X. Wen, Y. Liu, and M. Li, "Evenness-aware data collection for edge-assisted mobile crowdsensing in Internet of Vehicles," *IEEE Internet Things J.*, vol. 10, no. 1, pp. 1–16, Jan. 2023, doi: 10.1109/JIOT.2021.3095285.

[149] H. Kim and J. Ben-Othman, "Eco-friendly low resource security surveillance framework toward green AI digital twin," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 377–380, Jan. 2023, doi: 10.1109/LCOMM.2022.3218050.

[150] S. Zhu, K. Ota, and M. Dong, "Green AI for IIoT: Energy efficient intelligent edge computing for industrial Internet of Things," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 79–88, Mar. 2022, doi: 10.1109/TGCN.2021.3100622.

[151] H. Feng, Z. Cui, C. Han, J. Ning, and T. Yang, "Bidirectional green promotion of 6G and AI: Architecture, solutions, and platform," *IEEE Netw.*, vol. 35, no. 6, pp. 57–63, Nov. 2021, doi: 10.1109/MNET.101.2100285.

[152] H. Ryden, H. Farhadi, A. Palaios, L. Hevizi, D. Sandberg, and T. Kvernvik, "Next generation mobile networks' enablers: Machine learning-assisted mobility, traffic, and radio channel prediction," *IEEE Commun. Mag.*, early access, Jun. 19, 2023, doi: 10.1109/MCOM.001.2200592.

[153] Y. Ren, R. Xie, F. R. Yu, T. Huang, and Y. Liu, "Green intelligence networking for connected and autonomous vehicles in smart cities," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 3, pp. 1591–1603, Sep. 2022, doi: 10.1109/TGCN.2022.3148293.

[154] R. Mishra, H. P. Gupta, R. Kumar, and T. Dutta, "Leveraging augmented intelligence of things to enhance lifetime of UAV-enabled aerial networks," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 586–593, Jan. 2023, doi: 10.1109/TII.2022.3197410.

[155] G. Damigos, T. Lindgren, and G. Nikolakopoulos, "Toward 5G edge computing for enabling autonomous aerial vehicles," *IEEE Access*, vol. 11, pp. 3926–3941, 2023, doi: 10.1109/ACCESS.2023.3235067.

[156] G. Raja, S. G. Senthivel, S. Balaganesh, B. R. Rajakumar, V. Ravichandran, and M. Guizani, "MLB-IoD: Multi layered blockchain assisted 6G Internet of Drones ecosystem," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 2511–2520, Feb. 2023, doi: 10.1109/TVT.2022.3213567.

[157] A. M. H. Alibraheemi, M. N. Hindia, K. Dimyati, T. F. T. M. N. Izam, J. Yahaya, F. Qamar, and Z. H. Abdullah, "A survey of resource management in D2D communication for B5G networks," *IEEE Access*, vol. 11, pp. 7892–7923, 2023, doi: 10.1109/ACCESS.2023.3238799.

[158] A. Bani-Bakr, M. N. Hindia, K. Dimyati, Z. B. Zawawi, and T. F. T. M. N. Izam, "Caching and multicasting for fog radio access networks," *IEEE Access*, vol. 10, pp. 1823–1838, 2022, doi: 10.1109/ACCESS.2021.3137148.

[159] M. A. Maruf, A. Singh, A. Azim, and N. Auluck, "Faster fog computing based over-the-air vehicular updates: A transfer learning approach," *IEEE Trans. Services Comput.*, vol. 15, no. 6, pp. 3245–3259, Nov. 2022, doi: 10.1109/TSC.2021.3099897.

[160] A. Abouaomar, A. Taik, A. Filali, and S. Cherkaoui, "Federated deep reinforcement learning for open RAN slicing in 6G networks," *IEEE Commun. Mag.*, vol. 61, no. 2, pp. 126–132, Feb. 2023, doi: 10.1109/MCOM.007.2200555.

[161] M. A. Elsaid, "A novel data offloading with deep learning enabled cyberattack detection model for edge computing in 6G networks," in *AI-Enabled 6G Networks and Applications*. Hoboken, NJ, USA: Wiley, 2023, pp. 17–33.

[162] J. Du, B. Jiang, C. Jiang, Y. Shi, and Z. Han, "Gradient and channel aware dynamic scheduling for over-the-air computation in federated edge learning systems," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1035–1050, Apr. 2023, doi: 10.1109/jsac.2023.3242727.

[163] E.-K. Hong, I. Lee, B. Shim, Y.-C. Ko, S.-H. Kim, S. Pack, K. Lee, S. Kim, J.-H. Kim, Y. Shin, Y. Kim, and H. Jung, "6G R&D vision: Requirements and candidate technologies," *J. Commun. Netw.*, vol. 24, no. 2, pp. 232–245, Apr. 2022, doi: 10.23919/JCN.2022.000015.

[164] O. Chukhno, O. Galinina, S. Andreev, A. Molinaro, and A. Iera, "Interplay of user behavior, communication, and computing in immersive reality 6G applications," *IEEE Commun. Mag.*, vol. 60, no. 12, pp. 28–34, Dec. 2022, doi: 10.1109/MCOM.009.2200238.

[165] Z. Wei, H. Qu, Y. Wang, X. Yuan, H. Wu, Y. Du, K. Han, N. Zhang, and Z. Feng, "Integrated sensing and communication signals toward 5G-A and 6G: A survey," *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11068–11092, Jul. 2023, doi: 10.1109/JIOT.2023.3235618.

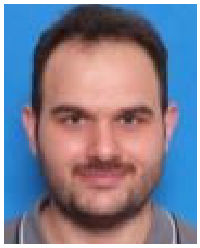
[166] U. Demirhan and A. Alkhateeb, "Integrated sensing and communication for 6G: Ten key machine learning roles," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 113–119, May 2023, doi: 10.1109/MCOM.006.2200480.

HAFIZ MUHAMMAD FAHAD NOMAN (Graduate Student Member, IEEE) received the B.E. degree in electronic engineering from COMSATS University, Islamabad, Pakistan, in 2012, and the M.E. degree in computer engineering from the University of Engineering and Technology (UET), Lahore, Pakistan, in 2016. He is currently pursuing the Ph.D. degree in electrical engineering (major in artificial intelligence and wireless communication) with the Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia. He has more than eight years of research and development experience in academia and industry. His research interests include machine learning, wireless networks, D2D communication, and energy-efficient 6G networks. He has been awarded the HEC Pakistan Overseas Scholarship for the Ph.D. studies, in 2022

EFFARIZA HANAFI (Member, IEEE) received the B.Eng. degree (Hons.) in telecommunications engineering from The University of Adelaide, Adelaide, Australia, in 2010, and the Ph.D. degree in electrical and electronic engineering from the University of Canterbury, Christchurch, New Zealand, in 2014. She joined the University of Malaya, Kuala Lumpur, Malaysia, where she is currently a Senior Lecturer of electrical engineering. She has authored or coauthored journal articles in ISI-indexed publications and refereed conference papers. Her research interests include multiple antennas systems, RFID, cognitive radio networks, cooperative communications in wireless communications, the Internet of Things, and 5G networks and beyond.

KAMARUL ARIFFIN NOORDIN (Senior Member, IEEE) received the B.Eng. (Hons.) and M.Eng. degrees from the University of Malaya, Kuala Lumpur, Malaysia, in 1998 and 2001, respectively, and the Ph.D. degree in communication systems from Lancaster University, U.K., in 2009. He is currently an Associate Professor with the Department of Electrical Engineering, University of Malaya. His research interests include resource allocation in wireless networks, cognitive radio networks, device-to-device communications, network modeling, and performance analysis.

KAHARUDIN DIMYATI (Member, IEEE) received the B.E. degree from the University of Malaya, Malaysia, in 1992, and the Ph.D. degree from the University of Wales, Swansea, U.K., in 1996. He is currently a Professor with the Department of Electrical Engineering, Faculty of Engineering, University of Malaya. Since joining the University of Malaya, he is actively involved in teaching, postgraduate supervision, research, and administration. To date, he has supervised 15 Ph.D. students and 32 master's students by research students. He has published more than 100 journal articles. He is also a member of IET and IEICE. He is also a Professional Engineer and a Chartered Engineer.



MHD NOUR HINDIA (Member, IEEE) received the Ph.D. degree from the Faculty of Engineering in Telecommunication, University of Malaya, Kuala Lumpur, Malaysia, in 2015. He is currently a Postdoctoral Fellow with the Faculty of Engineering in Telecommunication, University of Malaya. He is also involved with research in the field of wireless communications, especially in channel sounding, network planning, converge estimation, handover, scheduling, and quality of service enhancement for 5G networks. He has authored or coauthored a number of science citation index journals and conference papers. He has participated as a reviewer and a committee member of a number of ISI journals and conferences.

ATEF ABDRABOU (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada.

In 2010, he joined the Department of Electrical Engineering, United Arab Emirates University, Al-Ain, UAE, where he is currently an Associate Professor. His research interests include smart grid communication, vehicular communication, self-organizing wireless networks, radio resource allocation, the Internet of Things/machine-to-machine communication, and machine learning applications in wireless communication. He was a technical program committee member and the session chair at many IEEE international conferences. He is an Associate Editor of the IEEE Transactions on Mobile Computing.

FAIZAN QAMAR (Member, IEEE) received the B.E. degree in electronics from Hamdard University, Karachi, Pakistan, in 2010, the M.E. degree in telecommunication from NED University, Karachi, in 2013, and the Ph.D. degree in wireless networks from the Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia, in October 2019. He is currently a Senior Lecturer with the Faculty of Information Science and Technology (FTSM), Universiti Kebangsaan Malaysia (The National University of Malaysia—UKM), Selangor, Malaysia. He has more than nine years of research and teaching experience. He has authored or coauthored numerous ISI and Scopus journals and IEEE conference papers. His research interests include interference management, millimeter-wave communication, ad-hoc networks, machine learning, the Internet of Things, D2D communication, and quality of service enhancement for future wireless networks. He is serving as a reviewer in more than 15 high-reputation journals with various publishers, such as IEEE, Elsevier, Springer, Wiley, and Hindawi.