Received 31 July 2022, accepted 23 August 2022, date of publication 29 August 2022, date of current version 8 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3202956

SURVEY

AI-Enabled UAV Communications: Challenges and Future Directions

AMIRA O. HASHESH[1,2], SHERIEF HASHIMA[3,4], (Senior Member, IEEE), ROKAIA M. ZAKI[1,5], MOSTAFA M. FOUDA[6,7], (Senior Member, IEEE), KOHEI HATANO[3,8], AND ADLY S. TAG ELDIEN[1]

[1]Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University, Cairo 11672, Egypt
[2]High Institute for Engineering & Technology—Obour, Cairo 11828, Egypt
[3]Computational Learning Theory Team, RIKEN-Advanced Intelligence Project (AIP), Fukuoka 819-0395, Japan
[4]Engineering Department, Nuclear Research Center (NRC), Egyptian Atomic Energy Authority, Inshas, Cairo 13759, Egypt
[5]Higher Institute of Engineering and Technology, Kafr El-Shaikh 33514, Egypt
[6]Department of Electrical and Computer Engineering, College of Science and Engineering, Idaho State University, Pocatello, ID 83209, USA
[7]Center for Advanced Energy Studies (CAES), Idaho Falls, ID 83401, USA
[8]Faculty of Arts and Science, Kyushu University, Fukuoka 819-0395, Japan

Corresponding authors: Sherief Hashima (sherief.hashima@riken.jp) and Mostafa M. Fouda (mfouda@ieee.org)

This work was supported by JSPS KAKENHI Grant Numbers JP21K14162 and JP22H03649, Japan.

ABSTRACT Recently, unmanned aerial vehicles (UAVs) communications gained significant concentration as a talented technology for future wireless communications using its remarkable advantages and broad applicability. Furthermore, UAV networks' high complex configurations and designs encourage researchers to leverage relevant artificial intelligence (AI) techniques for better beyond fifth-generation (B5G)/sixth-generation (6G) services. This article summarizes AI-aided UAV solutions designated for forthcoming wireless networks. Besides, we deliver a comprehensive summary of machine learning (ML) approaches, including their applications and valuable contributions towards effective UAV network implementations, particularly advanced ML ones like bandits, federated learning (FL), meta-learning, etc. Finally, detailed UAV communication-related future research scopes and challenges is highlighted.

INDEX TERMS Unmanned aerial vehicles (UAVs), artificial intelligence (AI), deep learning (DL), meta-learning, federated learning (FL), reinforcement learning (RL).

I. INTRODUCTION

Beyond fifth-generation (B5G) & sixth-generation (6G) systems are primarily marked by ultra enormous connections, high-speed transmission rates (Gbps/Tbps), and low latency. Achieving these goals plus the quick expansion of the internet of things (IoT) implementations is challenging, especially in high dynamic and heterogeneous scenarios. An advantageous technique is via embracing unmanned aerial vehicles (UAVs) as flying clients or airborne base stations (BSs). Also, UAV-aided communications can enhance the network in emergency scenarios via delivering quick service retrieval and offloading in enormously overcrowded designs or even stricken areas [1]. Most remarkable practical UAV applications are shown in Figure 1. Still, the primary cons of UAVs are their limited battery life and poor processing power [2], [3], [4], [5], [6], [7], [8]. As in [9], UAVs can be considered the most appropriate candidate to enhance the performance and overcome the restrictions of ground networks, so it surveyed research issues in UAV communication system.

Furthermore, most existing commercial UAVs maneuver more than two hours before recharging their batteries. Besides that, complicated algorithms require high central processing unit (CPU) and graphics processing unit (GPU) capacities, which cannot operate on board due to the tight computing capacity of the UAV. Meanwhile, machine learning (ML) emerged as a sub-field of artificial intelligence (AI), which has become common in scientific study, introducing a novel approach known as the black-box strategy, which focuses solely on inputs and results. Additionally, the

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

vast amount of data available today and the availability of high-performance processors and strong GPUs aided the development of UAVs. More Precisely, ML is one of the trending domains in which machines are given intelligence and become intelligent to do tasks more efficiently than human beings. Therefore, nowadays, ML is being employed in various industries more than expected. Several sub-fields of AI have also emerged, including deep learning (DL), reinforcement learning (RL), federated learning (FL), etc., to handle different wireless communication difficulties, including UAV-related ones. DL utilizes layers of artificial perceptrons to simulate human thinking. It is significantly used in speech recognition, computer vision, and natural language processing. In addition, RL debuted in 1979 [10] and is a very active ML branch that has quickly evolved and matured, where an agent learns how to take proper actions to maximize his/her payoff. Hence, RL efficiently handles exploiting or exploring many available states. In contrast to DL, in robots, RL is commonly leveraged for path planning and learning complex tasks in robots. It's also used in a range of decision-making problems in which a goal-oriented agent interacts with a specific environment. Furthermore, FL, proposed by Google in 2016 [11], was created to facilitate data-decentralized network systems. It secures a highly centralized model training on devices that share decentralized data without transferring the data to a local shared unit to put it. Specifically, it runs various ML algorithms throughout decentralized data infrastructure.

Hence, leveraging AI for various UAV-related problems is a challenging and intelligent solution to such problems due to several AI merits. Although conventional methods succeeded in solving such issues, their solutions are still high complexity and consume much time with low accuracy. Hence,it is vital to investigate efficient AI-assisted solutions, especially self-decision-making ones, outperforming classical ones in complicated and powerful scenarios. Originally, UAVs were intended to be operated by people. However, with the rise of ML, it has become fashionable to sell smart UAVs. In this case, AI can execute distinct tasks using data collected by drone sensors. Besides, ML-assisted solutions can enhance the energy efficiency of UAVs through efficient resource management and interference mitigation. Furthermore, ML-aided-well planned trajectory planning helps in equipping UAVs with the proper battery capacity to avoid obstacles and plan their route autonomously, which supports more clients and prolongs battery life time. For example, "follow me" drones have recently become a huge hit in the market. These drones follow and film their owner with clever obstacle avoidance and target tracking algorithms, providing excellent video footage. Furthermore, vital UAV-related applications such as surveillance, traffic control, and landing site recognition are just a few applications that may be intelligently enhanced via distinct AI algorithms. In addition, leveraging existing state-of-the-art ML-based computer vision algorithms for picture enhancement approaches via UAVs applications is a promising direction. As a result, employing ML techniques

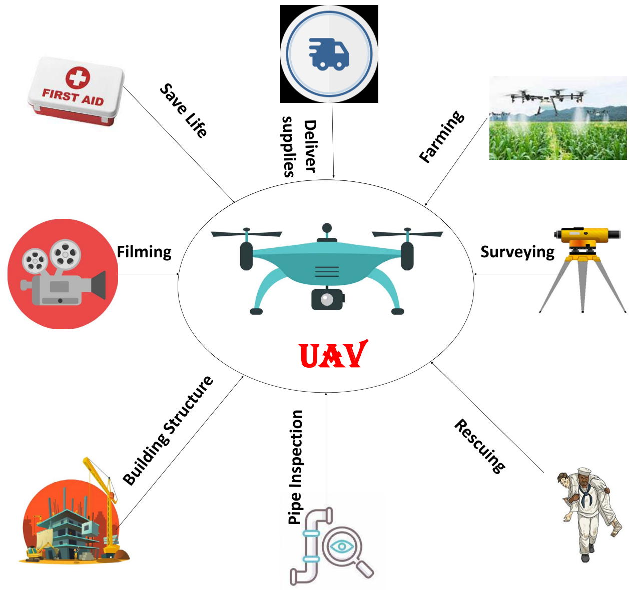


FIGURE 1. UAV promising applications.

to automate complicated UAV-related tasks and intelligently improve overall system efficiency can significantly increase the overall UAV's network performance.

We highlight most UAV-surveyed work and analyze them as shown in Table 1. Although there are related survey work [12], [13], [14], our paper is up to date and handles new related AI topics such as meta-learning and FL. Although the works in [15] and [16] reviewed RL and DL approaches for UAVs, they didn't directly address up-to-date AI methods for UAV applications, plus still some critical applications are missing. However, all of the previous papers did not directly address AI schemes for UAV applications. In [17], offered a complete overview of certain possible AI applications in UAV-based networks by dividing it into supervised and unsupervised approaches. Still, it has a restriction in that it gives a broad overview of all techniques and does not focus on individual lines. The work of [18] surveyed UAVs of different designs with different AI techniques, but it did not overview all methods. Furthermore, the work of [19] defined vital topics connected to UAVs and contemporary machine learning methods and presented a list of relevant courses and surveys. It examines flocks from the perspective of several open topics in which ML may be used to address various flock concerns. Motivated by the enormous importance of UAVs in future daily lives and the continuous development of AI schemes, we investigate different AI-assisted UAV communication approaches. Furthermore, future related challenges and possible solution scenarios are highlighted. This paper is organized as follows: Section II surveys Era before AI. Section III highlights different ML schemes and terminologies and their proper applications. Section IV goes over the difficulties that ML-based solutions have solved. Section V summarizes the future work, including different challenges. Finally, Section VI concludes the paper.

TABLE 1. Analysis AI-Enabled UAV communications.

Reference	Main Direction	Comparison	Future Challenge	Contribution
[13] ANN Only	Review ANN in general not specified with UAV.	✓	✓	Presented comprehensive tutorials on the use of ANNs-based ML for enabling various applications. It reviewed ANN-assisted UAV and talked about it briefly.
[15] DL only	Deep Learning Methods and Applications for UAV		✓	Reviewed reported uses and applications of deep learning schemes for UAVs, including the most relevant developments as well as their performances and limitations.
[16] ML	ML assisted Wireless UAV Networks		✓	Presented an specified overview of ML applications in aerial wireless networks with ML concepts and categories.
[14] ML	ML assisted UAV		✓	Presented a classification of ML techniques based on the communication and network aspects assisted-UAVs and reviewed different topics of UAVs.
[17] RL and FL Only	RL and FL assisted UAV		✓	Explored a research direction where ML techniques are used to enhance the performance of UAV networks, then reviewed RL and FL only with UAV applications.
[19] ML	ML assisted UAV Flocks	✓	✓	Surveyed several issues relating to UAV flocks formation, maintenance, and related challenges, it discussed from UAV flocks, not a single UAV scenario.
This paper	AI assisted UAV	✓	✓	Surveys all AI techniques utilized to enhance the performance of UAV and reviews researchers' effort to solve problems with these techniques, then shows the difference between them as reviewed them in categories depended on AI technique and showed which one is the most used. Finally, highlights hot research areas in UAV and our opinion in promising future techniques.

II. NON AI-BASED SOLUTIONS

Herein, we summarize different ordinary mathematical optimization solutions for UAV problems as shown in Table 2. In [20], the authors considered a single-cell scenario with many UAV-based users. To allow multi-UAV communications, they investigated two transmission modes called UAV to network and UAV to UAV. To enable UAV-to-X to communications, they designed a cooperative UAV sense and send protocol, then defined the subchannel allocation and UAV speed optimization issue for uplink sum-rate maximization. It maximized sum-rate uplink but in high complexity way. In [21], the researchers examined a multiuser communication design, where a single UAV base station (BS) with a single antenna might support many ground users via non-orthogonal multiple access (NOMA). Hence, such an optimization problem is a non-convex max-min rate specified by various parameters such as total power, total bandwidth, UAV altitude, transmit antenna beam-width, users' power, and bandwidth allocations. It optimized total power for system but with ideal assumptions. The authors of [22] looked at UAV-assisted wireless communication network systems with many energy-constrained ground terminals served by a UAV mounted access point (AP). They addressed two optimization challenges: a linear energy harvesting model and a realistic non-linear mode to optimize the ground terminals' minimal throughput. It optimized trajectory planning and resource allocation but with ideal assumptions. In [23], the authors jointly optimized both the UAV trajectory and NOMA precoding in a UAV-aided NOMA network scenario, where the UAV and BS simultaneously assist ground users (GUs). It Optimized the UAV trajectory and maximized the sum rate but with a long process time. In [24], the authors efficiently optimized UAV trajectory via decoupling state variables from timing variables. With the timing fixed, convex optimization may be used to maximize the state variables, and a nonlinear programming can be used to optimize the timing variables, resulting in a bi-level optimization problem. It Optimized trajectory planning but it neglected UAV battery consumption. In [25], the authors presented a UAV-aided NOMA scheme to achieve simultaneous wireless information and power transfer and guarantee the secure transmission for passive ground receivers, with nonlinear energy harvesting model, the throughput of passive receivers is maximized. However, it maximized the throughput using ideal settings. In [26], the authors proposed a new spectrum sharing scenario for a cognitive relay network. They investigated the optimization of the UAV relay's three-dimensional (3D) trajectory to improve the communication throughput performance of the secondary network subject to the interference constraints of the primary users (PUs). Still, they utilized ideal channel assumptions in their model. In [27], the authors suggested a 3D geometry-based stochastic multi-input multi-output (MIMO) channel model for intelligent reflecting surface (IRS)-aided UAV communications. The IRS is mounted outside the buildings to help the UAV transmitter reflect its signals to the ground level receiver and improve communication through passive beamforming. It looked at UBS consumers' spatial cross-correlation functions (CCFs) to see how IRS affected the underlying propagation channel. It more realistic IRS-assisted UAV design but with high complexity design, and lacks deep channel characteristics. In [28], the authors proposed a system in which UAVs are served as carriers of wireless power chargers to charge the energy-constrained devices to maximize the total amount of charging energy. Though, it introduced a higher charging amount delivery scenario with an ideal UAV's trajectory path. In [29], the authors proposed a low-complexity users BS (UBS)-assisted sleep strategy for the small-cell network to fill in the coverage holes as the low energy-efficiency BSs on the ground are switched off. It boosted the system's energy efficiency with long-time computation. In [30], the authors suggested a multi-UAV-assisted mobile edge computing (MEC) system to solve the computation efficiency maximization problem, which considers both computation bits and energy usage. They improved user allocation, power allocation, and UAV path planning via partial computation offloading mode. It proposed efficient computation offloading and trajectory scheduling for multi-UAV but ignored processing time. In [31], the authors presented a method to maximize energy efficiency for user equipment transmission, and the position of UAVs should be carefully evaluated to ensure a high quality of experience for user equipment with various priorities. It optimized UAV under several constraints but with ideal assumptions which are not practical. In [32], the authors suggested MEC as a UAV network where coalition leaders serve as servers to aid members with data computation. They looked into relative delay optimization in MEC-assisted UAV swarms. The computation offloading and channel access issues are jointly optimized according to the linked scheduling-resource allocation relationship.

III. ML SCHEMES AND TERMINOLOGIES

Various ML schemes have been discussed in several works of literature like [33], [34], [35], [36]. However, for transparency and absoluteness of the discussions, we shortly overview ML schemes and terminologies, including recent ones. The different learning schemes can be classified into one of the following varieties:

A. SUPERVISED-LEARNING (SL)

Here the model realizes a mapping formula, $y = f(x)$, employing documented data that delivers specimens of input-output ($x - y$) relation. Such a model anticipates forthcoming output (y_o) for a given trial input (x_o). The learning process is performed by evaluating the probability of the samples $p(y|x)$ [33], [35]. Regression and classification are the two primary SL-based predictive models. Regression frameworks employ statistical approaches to formulate the relation between descriptive variables and real-valued results. It anticipates the output utilizing either linear or sigmoid function approximations. However, classification models are broadly-used ML schemes that categorize data samples into one out of numerous available classes. Strictly speaking, it maps an input to one of the probable outputs. Ordinary classification models can be leveraged for UAV applications, such as support vector machine (SVM), K-nearest neighbor (KNN), and decision tree (DT) [37]. Besides, the current advanced GPUs designs permit extra sophisticated and artificial neural network (ANN) for huge-sized datasets. This Deep NN (DNN) design might incorporate convolutional neural network (CNN), Recurrent neural network (RNN), and Boltzmann machine have been used in many unexplored areas in wireless networks [34], [36]. We shortly emphasize the significant concepts of some widely-used categories as follows:

- **SVM**: is a binary-classification model that distinguishes between two distinct kinds of training samples [33].
- **KNN:** it is mainly utilized for both classification and regression tasks. At classification, we look for the K^{th} closest nearby training test sample x_o then classify concerning the bulk of the samples within the K^{th} most immediate neighbors. [33], [38]. Therefore, a particular subject is at $K = 1$, where x_o is designated to the class of the nearest sample using any specially selected distance function like the Minkowski norm L_p distance, etc. KNN evaluation relies on the weight of K, where fewer values conduct more precise classification outcomes with more noise sensitivity too. Also, more significant values decline noise sensitivity with less distinctive classes' accuracy.
- **DNN:** one of the primary usages of DNN is function approximation using weighted mixtures of accessible units (neurons) in a series of layers (input, hidden, and output ones). It is similar to brain processing with its complete details. Various DNN designs have been checked out in literature to reflect different functions, like mapping and regression [34], [36].
- **RNN:** is a subtype of NNs planned to formulate sequential data. The current network output is a function of the present input and the prior output via training memory cells. Checking distinct RNN architectures is promising for examining time series data in mobile systems plus their usage in speech recognition and NLP [34].
- **CNN:** as the name suggests, CNNs can automatically extract valuable features from raw input features more profitable than manual or human-based ones. It supposes locally connected filters rather than fully connected structures between layers to grab the spatial correlations [39]. CNN exploits two operations, namely convolution and pooling. The convolution uses multiple filters

TABLE 2. Mathematical optimization UAV solutions.

Reference	Scenario	ML Scheme	Objective	Limitations
[20] UAV-to-X Communications	Several UAVs send their collected data to BS	Iterative subchannel allocation and speed optimization algorithm (ISASOA)	Uplink sum-rate maximization	High complexity
[21] UAV+NOMA	UAVs as BS	Mathematical optimization algorithm	Optimize total power and BW	Ideal assumption and high complexity
[22] UAV-aided networks	UAV-mounted AP serves multiple ground terminals	Mathematical optimization algorithm	Optimize trajectory planning and resource allocation	Ideal setting
[23] UAV+ NOMA	cooperative UAV and BS to simultaneously serve GUs	Iterative optimization algorithm	Optimize the UAV trajectory and maximize sum rate	Long time computation
[24] UAV optimization	decoupling state variables from timing variables	Mathematical optimization algorithm	Optimize trajectory planning	Neglects UAV battery consumption
[25] UAV+NOMA	UAV-aided NOMA network	Iterative optimization algorithm	Maximize throughput of proposed system	Ideal scenario
[26] UAV based spectrum sharing	Joint 3D trajectory and resource optimization for a UAV relay-assisted CRN	Mathematical optimization schemes	Maximum throughput outage	Ideal channel assumption
[27] UAV+IRS	3D MIMO based Channel Model	Spatial CCFs	More realistic IRS assisted UAV design.	high complexity, lacks deep channel characteristics.
[28] UAV-Assisted IoT	Wireless Charging	MDP	Higher charging amount delivery	Ideal UAV's trajectory path
[29] UBS-enhanced small-cell Network	BS-assisted communication	Assisted sleep strategy	Boost the energy efficiency	Long Time computation
[30] UAV aided MEC	MEC Energy consumption	Iterative optimization algorithm	Propose efficient computation offloading and trajectory scheduling for multi-UAV	Ignores processing time
[31] Multi-UAV assisted MEC	Offloading setup to optimize energy	Heuristic Joint Power and Quality (HJPQ) algorithm	UAV optimization under several constrains	Ideal assumptions, not practical
[32] UAV swarms	MEC Scheduling urgent missions	Distributed offloading algorithm	Distributed UAV optimization coalitions	Ideal scenario

to extract features from the dataset and preserve their corresponding spatial information. Meanwhile, pooling (or sub-sampling) is leveraged to lower the dimensionality of the feature map via either max-pooling or average-pooling [34].

B. UNSUPERVISED-LEARNING

This procedure explores hidden patterns and configurations of the input data without data labels. Its main tasks are density estimation, clustering, and dimension reduction [33]. (1) Clustering means spine samples into sets or clusters, where at first each sample's class is unknown within the dataset. (2) Density estimation estimates the density of the data distribution in the feature space, revealing several essential features in the high-density areas, such as the Gaussian mixture model (GMM) technique. (3) Finally, dimension reduction, such as principal component analysis (PCA) and autoencoder, converts the data from a high-dimensional into low-dimensional space while preserving the data's primary arrangements. We summarize crucial unsupervised learning methods as follows:

- **K-means:** is a simple clustering technique that locates *m* sample optimal points in the feature space of m clusters. Each sample is designated to one cluster concerning the distance between individual points and representatives. Still, picking the optimal m-points is an NP-hard problem that can be estimated by operating a less complex iterative scheme via erratically determining initial *m*-points and allocating all samples to the initial k-points. After that, we obtain the mean per cluster and replicate the process until convergence occurs [40].
- **GMM:** belongs to density estimation methods. Its main target is to fit the data into a mixture (weighted linear combination) of k Gaussian probability distributions to address complex cluster arrangements. k manages the complexity of GMM, where incrementing k permits GMM to match any continuous distribution accurately. However, the larger k, the greater the risk of overfitting and the time required to estimate the mixing parameters using the log-likelihood method [41].

C. REINFORCEMENT-LEARNING (RL)

encompasses a variety of human-like learning processes based on trial-and-error. To achieve long-term gains, an RL player is rewarded or penalized for his actions. The agent receives recursive environmental feedback to assist in determining the appropriate actions at each step by following a policy that translates agent behavior from state to action. With uncertainty in the environment, a Markov decision process (MDP) may be used to describe the system's dynamics and maximize the objectives. [35].

- **Q-Learning:** is a prototype reinforcement learning strategy in which the agent does not need to know or have a model of the environment. The agent calculates and stores a Q-value for each stateaction pair in the Q-table from training. The Q-value is a long-term payoff. However, it is not ideal for big-scale issues because tables get too vast when the problems become more complex [42].

D. SELF LEARNING TECHNIQUES

1) MULTI-ARMED BANDITS (MABs)

The MAB problem is one of the sequence allocating. The player/learner tries to get the highest payout from a series of slot machine arms, where the payouts are distributed randomly. The title MAB arises from the notion of a gambler playing a group of slot machines inside a casino. He should decide, which machine/arm to play, the number of times, and the rank to play each. This is to determine whether to continue with a slot machine or change to another one. The tradeoff means the balance between exploiting the selection that gave the highest payoffs, current knowledge (the best-selected arm), and exploring new arms (unselected or rarely chosen elements,i.e, unknown environment) that might provide higher future profits. Although the old study of bandit problems since the 1930s, the exploration-exploitation conflict appears in several modern applications, like advertising, website optimization, resource allocation, and network routing [43]. Due to its merits, MAB algorithms have been used for different wireless communication problems such as D2D communications [44], [45], [46], [47], WSNs [48], Relay probing [49], IRS [50], hybrid band communications [51], [52], [53], mmWave beamforming [54], [55], [56], UAV communications [57], [58], etc. Generally, MAB can be divided into the following categories:-

- **Single Player MAB:** Within a restricted number of trials, a single player seeks to locate and pick the largest long-term reward arm [43]. First, the player gathers information about each slot machine (exploration) by inspecting a variety of accessible arms and finishing with the arm that pays the most. As a result, the player strives to strike a balance between playing with the arm with the highest possible payout thus far, i.e. exploitation, and exploring other arms, i.e. exploration. The player can precisely forecast each arm's due reward over a longer horizon time (investigation term). According to the allocation of awards, the MAB problem is stochastic or adversarial [43]. In stochastic bandits, the rewards of each arm are pulled independently (i.i.d), from unknown distributions to the players. Upper confidence bound (UCB), TS are the top most stochastic MAB algorithms [45]. In adversarial MAB, on the other hand, the rewards are determined by the hostile environment like in ϵ greedy, Exponential-weight algorithm for exploration and exploitation (EXP3), and EXP4 algorithms [43].
- **Multi-Player (MP)-MAB:** All players act in sequential trials simultaneously to obtain an anonymous reward [58]. If more than one player picks a similar arm, collisions happen. Later, players might distribute the rewards or disregard them upon the collision rule. Upon the mutual information among the players, multiplayer MAB schemes are classified as centralized and decentralized. In decentralized setup, each player selfishly plays his future trials based on his collected reward remarks without data interchange with other players [43], [58]. In the centralized model, though, the game is run collectively by exchanging complete findings. Compared to their centralized counterparts, collisions are unavoidable in a decentralized configuration. As a result, each player acts selfishly to investigate collisions and tries to bypass them during interaction with the environment to increase his profit.
- **Contextual Bandits:** Here, the player gains his awards from taking actions (selecting arms) over a sequence of trials considering side information about each arm called context [59], [60]. Hence, within each trial: 1) The player acts based on the current round's context (feature vector) and the previously earned prizes. 2) The player is

solely aware of the prize for the chosen arm. Contextual MAB (CMAB) is used in various essential applications, including online recommendations, mobile health applications, and clinical studies. Exploration is necessary to improve learning performance, whereas supervised ML provides the features needed to encode context. As a result, CMABs are the standard compromise between supervised learning and RL. The CMAB problem is usually handled by providing a linear relationship between the created reward and the circumstances in which it occurs like in LinUCB algorithm [43].

- **Sleeping Bandits:** Here, the action set is time-varying i.e., non stationary. Hence, at every round, both the reward process and the arm availability might be modeled as Markovian, adversarial or stochastic. Some arms are out from the game during the rounds. Hence, the performance loss is w.r.t. the best action as a benchmark which might not exist in some rounds [61].
- **Combinatorial Bandits:** It is a multi-variable bandit game where the player selects values from a group of variables [62], not from a single one. The number of probable selections per iteration is exponential in the number of discrete variables. Its applications include shortest-path problems, ranking, and multitask bandits.
- **Cascaded Bandits:** Here, the learner investigates a recommended list from the first to the last item with unknown attraction probabilities and then chooses the first attractive one [63]. The beforehand things are nonattractive, and the proceeding ones are still unobserved. The optimal list is the one with K items with a maximum probability of attractive items. At time t, the learner recommends to the client a list of K items out of total L ones and then observes the item's index that the client selects. If the user selects an item, the learner receives a reward of one. The learner's target is to maximize his cumulative payoff or minimize his total loss/regret concerning the list of K's most precious items.

2) FEDERATED LEARNING

FL allows its models to learn from various data sources across several places (e.g., local data centers, a central server) without sharing any training data. This permits personal data to be stored locally, lowering the risk of personal data breaches. There are two phases in ML: training and inference. Local ML models are trained on local heterogeneous datasets during training. Users of an ML application, for example, can discover errors in the ML program's predictions and rectify them. Local training datasets are created in each user's device due to this. The variables of the models are then regularly shared between these regional data centers. Many models encrypt these parameters before sending them. Data samples from the local area are not shared. This increases data security and protection. A worldwide model has been developed. Finally, the global model's properties are shared with local data centers so that they may incorporate the global model into their ML local models. A model is kept on the user device during inference so that predictions may be made fast utilizing the model on the user device [64]. In [65], a joint algorithm of UAV placement, power control, transmission time, model accuracy, bandwidth allocation, and computing resources, namely energy-efficient FL (E2FL) has been proposed, aiming to minimize the total energy consumption of the aerial server and users.

3) META-LEARNING

It is an ML sub-field known as "learning to learn." It is used to enhance the outcomes and performance of a learning algorithm by modifying specific components of the algorithm depending on the results of testing experiments. Here, all of the training dataset, the learning methodology, and the algorithm's parameters affect the learning model's performance. Hence, this necessitates a large number of tests. Meta learning methods help speed up the learning process, where better forecasts are made in less time. Researchers may use meta-learning to determine which algorithms produce the best predictions from datasets. Learning algorithms' information/foreknowledge is used as input to meta-learning algorithms. Then, they have predictions and offer data regarding the performance of these learning algorithms as an output. Metadata is data about data for non-technical consumers, such as size, resolution, style, date generated, and owner of a picture in a learning model. As a conclusion, meta-learning means learning new activities more quickly by using metadata [66].

4) TRANSFER LEARNING

Its goal is to help target learners enhance their performance on target domains by transferring information from several but related source domains. The need for a significant amount of target-domain data to generate target learners can be decreased. Transfer learning has become a prominent and promising field in machine learning due to its wide range of applications. Domain adaptation is modifying one or more source domains to transfer information and improve the target learner's performance. The domain adaptation method, which tries to narrow the gap across disciplines, is frequently used in transfer learning [67].

5) ADAPTIVE LEARNING

Here, the consequence of a decision is frequently unknown, and the effects might fluctuate over time. If choice results reflect a usual range of outcomes or signify a shift in the reward environment, they should have a significant impact on behavior and learning. As a result, practical learning and decision-making need the capacity to assess both expected and unexpected uncertainty (connected to the variability of findings) (associated with the variability of the environment). Understanding the computational and neurological basis and impacts of these two forms of luck and the interconnections between them is critical for understanding adaptive learning [68].

IV. ML-BASED SOLUTIONS

A. SL SOLUTIONS

Table 3 presents K-means and DL UAV solutions that solve several problems in UAV applications. In [69], the authors used NOMA to investigate a UAV-assisted VLC and constructed a combination issue of power allocation and UAV placement to optimize the total rate of all users, subject to limitations on power allocation, and user quality of service, and UAV position. It maximized the sum rate of all NOMA users but with a fixed UAV assumption. In [70], the authors proposed a distributed algorithm that allows UAVs to dynamically learn their optimal 3D locations and associate with ground users while maximizing the network's sum rate. When compared to both a centralized sub-optimal solution and a distributed approach based on the closest UAV association, the network's sum rate is improved but does not save the power of the system. [71], the authors presented a case study to demonstrate the effectiveness of intelligent UAV-assisted vehicular edge computing (VEC) architecture, a smart UAV-assisted VEC system envisioned to satisfy 6G Vehicle to Everything (V2X) requirements and provide 3D adaptive service coverage. In [72], the authors proposed a blockchain and AI-empowered telesurgery system towards 6G, which is a self-manageable, secure, transparent, and trustable system with massive Ultra-Reliable Low-Latency Communication (uRLLC). In [73], a novel UAV aerial video dataset (ManipalUAVid) is introduced for semantic segmentation. On the ManipalUAVid dataset, the performance of four semantic segmentation techniques is evaluated: conditional random Field, U-Net, Fully convolutional network, and DeepLabV3+. It introduced more clear shots, but it neglected processing time. In [74], with reduced model sizes and quicker computing speed, a novel lightweight AMC (LightAMC) technique is developed, which introduces a scaling factor for each neuron in a convolutional neural network (CNN) and enforces the sparsity of scaling factors using compressed sensing. It reduced model sizes and accelerated computation but with ideal assumptions.

DL solutions In [75], the authors offer an approach that uses semi-supervised techniques to categorize an unlabeled training set that is utilized for training a CNN using multiple training strategies, as the number of labeled samples available to train the classifier decreases in contrast to the amount of unlabeled data. In [76], the authors suggest an effective alternative technique for allowing the UAV to independently establish its location without relying on the global positioning system (GPS) or sending messages.

DRL Based Solutions: Table 4 shows different solutions for UAV applications using DRL. In [77], the authors investigated the difficulty of designing a 3D UAV trajectory and band allocation that considers the UAV's energy usage and the fairness among ground users. First, model a quadrotor UAV's energy consumption as a function of its 3D mobility. The fair throughput is then defined and maximized within restricted energy, depending on the fairness and total throughput. It automatically adjusted of UAV's Flight speed but it worked with a single UAV. The work in [78], investigated the cellular networks using UAVs, in which a UAV operates as a flying relay to unload a portion of the data flow from one congested cell to another. It used a plausible air-to-ground channel model and a practical geographical distribution of data traffic. The quality of service is described as a UAV utility function based on a packet loss ratio-related consumers' cost function to indicate the UAV's performance improvements. To optimize the UAV utility function, a joint optimization problem must be solved. It succeeded in maximizing the throughput, but it did not consider all users that used the system. Furtheremore, the framework in [80] investigated the robust and secure transmission for Reconfigurable intelligent surface (RIS)-aided mmWave UAV communications. It proposes an algorithm to effectively tackle the concerning issues by maximizing the sum secrecy rate of all legitimate users. It gets results by combining UAV trajectory optimization and active (passive) beamforming. A better performance can be achieved compared to a variety of benchmarks. It improved the sum secrecy rate of the system but with neglecting processing time computations. In [79], due to complicated limitations, it suggested a UAV trajectory planning model for data collection intending to minimize expired data packets across the sensor system and then relaxed the cryptic original issue into a min-max-age of information (AoI)-optimal route scheme. It solved the UAV path planning with unknown channel states but with specific area. In [81], the authors studied the topic of providing the optimum quality of service (QoS) in UAV-assisted cellular networks. To effectively optimize the usefulness of the UAV, it has suggested a combination design of access point selection and UAV path planning. It has presented a DRL-based method to teach the UAV to seek places with superior channel states and a game theory-based access point selection algorithm to allow users to select the correct access point autonomously based on the cost function. It minimized the content delivery delay but battery life time remained short. In [82], the authors investigated The The cache-enabling UAV NOMA networks,, which UAV base stations aid, and are designed for a mix of augmented reality and traditional multimedia applications. DRL optimizes user association, NOMA power allocation, UAV deployment, and UAV caching placement altogether to reduce content delivery time. It controlled continuous action space but with single agent. In [83], it proposed a UAV-aided MEC framework, as several UAVs. with varying trajectories fly over the target region and assist the ground based user equipment. By optimizing each UAV's trajectory and offloading decision from all the user equipment, a multi-agent DRL-based trajectory control algorithm can jointly maximize the fairness among all the user equipment and the fairness of user equipment-load of each UAV, as well as minimize the energy consumption of all the user equipment. It managing the trajectory of each UAV independently but it did not take cooperative decision.

TABLE 3. k-means and DL-based UAV solutions.

Reference	Scenario	ML Scheme	Objective	Limitations
[69] UAV-assisted VLC	Sum-Rate Maximization of all NOMA users	HHO+ANN	Maximize sum rate by the proposed algorithm/HHO trainer	Fixed UAV assumption
[70] Multi-UAVs networks	Design UAV network	k-means (SL)	Network's sum-rate is improved	Saving power
[71] UAV assisted VEC	UAV-assisted VEC for 6G IoV networks	AI	Reviews UAV with edge computing in 6G	Limited review
[72] Drone-assisted remote Surgery System	6G	AI- XGBoost algorithm	Integrate AI techniques to predict the type of disease and surgery	Speed to get solution
[73] New UAV aerial video dataset	Video dataset modelling	DL	Present a new UAV aerial video datase for semantic segmentation	Processing time
[74] Light weight automatic modulation classification	Automatic modulation classification	CNN	LightAMC method reduce model sizes and accelerate computation with the slight performance loss	Ideal assumption

In [84], the topic of reducing the normalized weighted sum of AoI for a UAV-assisted wireless network in which a UAV collects status update packets from energy-constrained ground nodes was discussed by the authors. The problem was first started as a mixed-integer program. It then suggested a convex optimization-based technique for obtaining the UAV's ideal flight trajectory and time instants on updates for a given scheduling strategy. It optimized the UAV's flight trajectory and minimized the normalized weighted sum of AoI but it used single agent. In [85], the authors looked at how to establish UAV-assisted MEC networks in a short amount of time while simultaneously serving several users. It also presented an end-to-end DRL model to learn and optimize task offloading and UAV trajectory control. The proposed approach optimizes many criteria, including computing delay and energy consumption of the UAV-assisted MEC network, by controlling the fraction of offloading jobs and UAV trajectory. It optimized the offloading task ratio and minimized the overall energy consumption in UAV but used single agent. In [86], the authors investigated the usage of UAVs to assist intelligent transportation system applications. It looked at the topic of minimizing the predicted weighted sum AoI of cars in a vehicular network by optimizing the trajectory of various UAVs and scheduling policies. It minimized the expected weighted sum AoI but did not consider power consumption. In [87], the authors suggested a cloud-assisted joint charging scheduling and energy management framework for UAVs, and then used multi-agent DRL to design and implement cooperative energy sharing across towers, resulting in intelligent energy sharing. It can be seen that the two approaches are linked and that they should be controlled, coordinated, and harmonized by a centralized orchestration manager, with fairness, energy efficiency, and cost effectiveness in mind. In [88], the authors looked at multi-dimensional resource management for vehicular networks using UAVs. The macro eNodeB and UAV, both mounted with MEC servers, work together to make association choices and assign appropriate resources to vehicles to enable on-demand resource access effectively. It formulated the resource allocation at the MEC servers as a distributive optimization problem to maximize the number of offloaded tasks while satisfying their heterogeneous QoS requirements and then solved it with a probabilistic multi-agent deep deterministic policy gradient (PMADDPG) based method because there is no central controller. In [89], as a MEC framework with a renewable power supply, the researchers devised a UAV-assisted compute offloading technique. The suggested model considers energy arrival instability, stochastic computing demands, and a changing channel state. Due to the state's complexity, UAV-assisted computed offloading for MEC based on DRL was proposed to reduce the overall cost, which is the weighted sum of delay, energy consumption, and bandwidth cost. In [90], the authors provided a space-air-ground integrated network edge/cloud computing design for offloading computation-intensive applications even considering remote energy and computation restrictions, where flying UAVs provide near-user edge computing and satellites provide cloud computing access. In [91], to determine the best solution for energy-harvesting time scheduling in UAV-assisted device To device (D2D) communications, the authors suggested a unique model based on DRL. The UAV is considered to fly around a central point to make the system model more realistic. The D2D users move in a continuous random walk. The channel state information encountered during each time slot is randomly time-variant.In [92], the authors presented a UAV system that uses wireless energy transfer to collect data from various geographical regions and deliver it to its destination modeled mobility, energy storage, and data storage patterns to account for time-variant system states detected by the UAV and their effects on decision-making. In [93], for the air-ground coordinated communications system, the authors suggested aerial to ground (A2G)-PMADDPG. By coordinating both UAV-BSs and GUs, the proposed algorithm allows UAV-BSs to offer equitable communication services for GUs on the ground. Each GU maximizes its throughput by selecting the appropriate UAV-BS to access, and each GU maximizes the fair throughput by designing a trajectory. Simulation results show that the approach outperforms existing benchmarks regarding fairness index, total throughput, and minimum throughput. A NOMA-based UAV-assisted network is gaining traction as a viable solution for overcoming various Like For 5G and B5G wireless networks, high spectrum

TABLE 4. DRL-based UAV solutions.

Reference	Scenario	ML Scheme	Objective	Limitations
[77] UAV trajectory design	3D UAV trajectory design and band allocation	DRL	Automatic adjustment of UAV's Flight speed and direction plus out performing the baseline methods.	Single UAV only
[78] Air-Ground Coordinated communications system	Trajectory design	Multi-Agent DRL	Enable GUs to maximize their own throughput and UAV-BSs to provide fair and high throughput communication service	all users not considered
[79] UAV-Assisted cellular networks	Design path planning and access point selection	DRL	Solve the UAV path planning subproblem in an area with unknown channel states	Specific area
[80] Re-configurable intelligent surface aided millimeter wave UAV communications	RIS elements and the UAV trajectory are jointly designed	DRL	Improve the sum secrecy rate of system	processing time not considered
[81] UAV+NOMA networks	Jointly optimize user association, power allocation of NOMA and UAVs placement	DRL	Minimize the content delivery delay. Mitigate the unobservable interference from the decision	Battery life time
[82] Multi-UAV-assisted IoT networks	MEC allocates resources	DRL	Control continuous action space and large dimensional state space. Minimize the computation costs	Single agent
[83] UAV assisted MEC	Trajectory optimize the geographical fairness	Multi-agent DRL	Managing the trajectory of each UAV independently	Not cooperative decision
[84] UAV-assisted networks	Battery-constrains trajectory	DRL	Optimize the UAV's flight trajectory and minimize the normalized weighted sum of AoI	Single agent
[85] UAV-assisted MEC	MEC trajectory control	DRL	Optimize the offloading task ratio and UAV trajectory to minimize the overall energy consumption in UAV	Single agent
[86] UAVs in intelligent transportation systems	Jointly optimize the trajectories of UAVs and scheduling policies	DRL	Finding the trajectories of the deployed UAVs and scheduling of status-updates to minimize expected weighted sum AoI	Power consumption
[87] Cloud-assisted multi-UAV	Cloud-assisted joint charging scheduling energy management	Multi-agent DRL	Present a new cooperative energy management system in edges/charging towers	Time to Charge
[88] UAV-assisted vehicular networks	Multi-access edge computing Resource management	Multi-agent DRL	Propose resource management technique	Energy computation
[89] UAV-assisted computation offloading	MEC design an UAV-assisted computation offloading scheme	DRL, K-means	Design an UAV-assisted computation offloading scheme with renewable power supply	Path planning
[90] Space/Aerial-assisted computing offloading and IoT	A space-air-ground integrated network edge/cloud computing architecture considering remote energy and computation constraints	DRL	Integrate the historical network information to learn the system dynamics. Adopt network virtualization to flexibly allocate the resources of the edge server	Time factor
[91] UAV assisted D2D communications	Optimal solution for Energy Harvesting time scheduling	DRL	Capability of solving real time allocation problems	Battery life time
[92] mMTC and IoT	UAV-assisted wireless energy and data transfer	DRL	The proposed MDP assisted data delivery and energy charging scheme outperforms conventional techniques.	Limited Battery capacity
[93] Intelligent trajectory design for UAV	Trajectory MIMO-UAV	DRL	Maximizing the average system secrecy rate of the system under some conditions	Power consumption
[94] Wireless-Powered UAV Networks	Trajectory and energy of UAV	DRL	Maximizing the sum-energy received by all UAVs and optimizing the energy loading process	Ideal assumption
[95] Synchronizing UAV Teams	Timely data collection and energy transfer for UAV	DRL	Maximize the throughput of IoT devices, minimize the energy utilization of UAVs, and enhance the energy transfer.	Ideal assumption
[96] Trajectory Optimization for Air-ground Cooperative Emergency Networks	Trajectory of UAV	Federated-DRL	Reduce the communication overhead with a distributed architecture.	Battery life time

efficiency and enormous connections are required, especially when IoT devices are placed in a disaster region. It maximized the average system secrecy rate but did not calculate power consumption. In [96], the authors proposed federated multi-agent deep deterministic policy gradient (F-MADDPG) based trajectory optimization algorithm to maximize the average spectrum efficiency. Because of difficulty coordinate between UAVs to improve the performance in terms of Wireless Energy Transfer (WET) and Wireless Information Transmission (WIT), authors of [95], proposed a Multi-Agent Deep Reinforcement Learning (MADRL) method, called TEAM to divide UAVs into two teams to behave as data collectors and energy transmitters to maximize the throughput of IoT devices, minimize the energy utilization of UAVs, and enhance the energy transfer.In [94], the authors leveraged MADRL method to optimize the task of energy transfer between Flying Energy Sources (FESs) and UAVs to maximize the sum-energy received by all UAVs, optimize the energy loading process and compute the most energy-efficient trajectories.

B. USL SOLUTIONS

In [97], small cell networks (SCNs) provide a cost-effective coverage option for high-data-rate wireless applications. However, with SCNs, appropriate management of backhaul lines to small cell BSs is a difficult task (SCBSs). Therefore, researchers use the notion of using UAVs to offer a connection between SCBSs and the core network to construct a solid backhaul link, where perfect line-of-sight (LoS) communication between the SCBSs and the core network plays a critical role. We examine the relationship between SCBSs and UAVs by considering a variety of communication-related aspects, such as the data rate limit and backhaul bandwidth resources. In [98], the authors planned predictive models with one-class support vector machines (OC-SVM) and K-means clustering to detect eavesdropping attacks. They also propose a framework for creating features of testing data from wireless signals and another framework for generating training data to prepare datasets for training predictive models.

C. RL SOLUTIONS

In Table 5, we focus on RL-based UAV solutions that solved several problems in UAV applications. In [99], the authors concentrated on a UAV-assisted wireless network where users can be scheduled to receive the uplink transmission from either an aerial or a terrestrial base station. The average long-term transmit power required by the users was reduced by dynamically optimizing user association and power allocation in each time slot. It enhanced power allocation and user association using UAV. but with fixed resource allocation. In [100], the authors considered the problem of content delivery to vehicles on road segments with either overloaded or no available communication infrastructure, resorting to tools such as proximal policy optimization, along with a set of crafted algorithms to solve our problem. It delivered high-bandwidth contents robustly but with short battery life time. In [101], the authors proposed Q-learning- based adaptive geographic routing to improve the converging speed and resource utilization of the geographic routing approaches in vehicular ad hoc networks (VANET). Autonomous vehicles (AVs) are deployed to guide the global transmission path and a Q-learning algorithm is exploited to help each node choose the best next hop in a specific area. In [102], the researchers looked at using UAV-assisted edge caching to help terrestrial vehicle networks transmit high bandwidth content files. It created a combination caching and trajectory optimization issue to judge content location, content distribution, and UAV trajectory to improve total network throughput. Due to complex constraints, it chose the optimal path scheme but did not consider saving power. In [104], The authors suggested an online RL UAV-assisted wireless caching system that optimizes the UAV trajectory, transmission power, and caching content scheduling all at the same time. It used the notion of request queues in wireless caching networks to define the combined optimization of online UAV trajectory and caching content delivery as an infinite-horizon ergodic to produce a QoS-optimal solution. It achieved online optimization of UAV trajectory but it did not calculate time consumption. In [105], for delay-tolerant wireless sensor network (WSN) applications, the authors suggested an autonomous UAV-based data collection system. The goal is to use a self-trained UAV as a flying mobile unit to gather data from ground sensor nodes geographically spread over a particular geographical area during a predetermined period. In [106], a UAV-assisted computation offloading model was developed by the authors, in which a group of UAVs flies about while offering value-added edge computing services. Multi-agent RL algorithms offered the target helper for the next task execution and the proportion of bandwidth allotted to communication, where two agents choose the target helper and bandwidth allocation. In [107], the authors recommended that several UAVs' paths be designed based on users' mobility data forecast. Combining trajectory design and power control challenge maximized the instantaneous total transmit rate while meeting customers' rate requirements. The authors of [108] investigated cache-enabled UAV cellular networks with NOMA support for colossal access. A mobile UAV BS, which caches some popular contents for wireless backhaul connection traffic unloading, assists in transmitting of a high number of multimedia material for ground users. In [109], the authors developed an onboard deep Q-network to reduce total data packet loss of sensing devices in UAV scenarios. In [110], the authors created an RL issue by modeling the motion-trajectory as MDP and using the UAV as the learning agent. It then proposed a pair of novel trajectory optimization algorithms based on stochastic modeling and reinforcement learning, which allowed the UAV to optimize its flight trajectory without requiring system identification. In [111], the authors used for cooperative search and rescue, UAVs and unmanned surface vehicles constitute a cognitive mobile computer network, where RL is utilized to design search paths and increase communication throughput.

TABLE 5. RL-based UAV solutions.

Reference	Scenario	ML Scheme	Objective	Limitations
[99] UAV assisted Communications	minimization of the users' long-term average consumed transmitted power	Relative Value Iteration (RVI) algorithm	Enhance power allocation and user association using UAV.	Fixed resource allocation
[100] UAV aerial-assisted vehicular networks	Trajectory selection	RL	Deliver high-bandwidth contents robustly	Battery life time
[101] UAV-assisted geographic routing	Vehicular ad hoc network	Adaptive-QL	Propose adaptive UAV routing technique	Network type
[102] UAV-assisted Data Sensing	Trajectory design	RL	Optimal path scheme due to complex constraints	Saving power
[103] UAVs-Enabled wireless networks	Like survey	FDL	Highlight applications of FDL in UAVs	Limited review
[104] UAV wireless networks with content and energy recharging	Trajectory planning	Online RL	Online optimization of UAV trajectory and radio resource with energy and content recharging with reduced-Complexity optimality conditions	Time consumption
[105] UAV-assisted data Collection for Wireless Sensor networks	Trajectory design	RL and QL	Enable autonomous navigation in an obstacle-constrained environment. Maintain the UAV safety from crashes due to energy depletion	Energy computation
[106] UAVs-assisted edge computing	Design computation offloading	Multi-Agent RL	Contrive the computation offloading problem to learn the near-optimal offloading policy by interactions with the environment	Battery life time
[107] Multi-UAV assisted wireless networks	Trajectory design power control	Multi-agent QL	Formulate on throughput maximization problem by designing the trajectory and power control of multiple UAVs	Time computation
[108] UAV NOMA networks	NOMA resource allocation	QL	Propose a framework of cache-enabling UAV NOMA cellular networks for content delivery of ground users in a hotspot area	Power consumption
[109] UAV-assisted online power transfer	Microwave power transfer minimize the overall data packet loss data allocation device charging	QL	Minimize the overall data packet loss of the sensing devices	Processing time
[110] Adaptive UAV-trajectory optimization	Maximizing the cumulative collected data Modelling the motion-trajectory	RL	Maximize the cumulative data volume of the UAV collected from the sensors	Energy calculations
[111] Group mobile computing for UAVs and USVs	UAVs as cognitive mobile computing network to improve communication throughput	Multi-agent RL	UAVs and USVs are jointed and model the path planning problem	Base station power
[1] Multiple-UAV networks deployment and movement design	Maximizing the sum mean opinion score of the users	Q-learning	Solve the problem of maximizing the sum the mean opinion score of the users	Power consumption
[112] Mobile UAV base stations	Beam selection	RL	Modeling of the beam selection of millimeter Wave base stations	Processing time
[113] NOMA-Based UAV-aided networks for emergency communication	Disaster NOMA energy efficient maximization	DQL	Throughput maximization for multi-UAV enabled NOMA networks	Fixed UAV

The authors of [1] proposed the challenge of joint non-convex 3D deployment and dynamic movement of UAVs to maximize ground users' total mean opinion score in quality of experience-driven deployment and dynamic movement of numerous UAVs. In [112], the 3D UAV aided mmWAve model was investigated to simulate beam selection and environmental responsiveness and regularly get near optimal evaluations by learning from current circumstances. In [113], the authors developed a NOMA-based UAV-assisted networks emergency communications architecture, in which catastrophe situations are split into three major categories: emergency areas, large regions, and dense areas. In disaster regions, a UAV outfitted with an antenna array might offer wireless coverage to several densely scattered devices.

D. MAB-BASED SOLUTIONS

Herein, we focus on the UAV MAB-based solution, as shown in Table 6. In [114], the authors looked at how a rotary-wing UAV can function as a wireless base station for emergency communication in a post-disaster environment with an uncertain user distribution. The goal of the described optimization challenge is to determine the ideal path that starts and ends at the same spot to serve as many people as feasible with limited battery capacity. This issue was reformulated using two extended MAB-enabled path planning algorithms. UAVs have been used as a critical alternative for ground communications in catastrophe zones such as earthquakes and exposed forests because of their benefits. In [115], the authors looked at radio resource allocation for a post-disaster surveillance system built with a cognitive radio network (CRN). They solved it with a dynamic spectrum access system and MAB. It maximized the total system rate but with ideal setting assumption. In [58], the difficulty of choosing a gateway UAV is solved. The major goal is to maximize the UAV relays' long-term average data rates while lowering the flight's battery cost, utilizing mmWave backhauling, which uses the 30 300 GHz band and antenna beamforming. In [116], the authors offered the MAB solution to improve the performance of any mobile networked device. Their results also showed that the 3D method optimizes technical resources compared to current single and 2-Dimensional algorithms, resulting in near to ideal performance throughout the average duration through machine learning of actual UAV communication settings. In [117], the authors used power control in combination with channel selection to examine anti-jamming vehicle to vehicle (V2V) communication in connected and autonomous vehicle (CAV) networks. The overall framework of cognitive risk control (CRC) is well tailored to assess and address the jamming problem by bringing a brain-inspired research instrument, cognitive dynamic system (CDS). Power control is specifically carried out via RL, with the results being assessed by a task-switch control module. The MAB issue is constructed based on the risk assessment when performing the channel-selection procedure. Still, their studied structure is simplified and needs to be more practical.

In [118], the plan target could be a secondary user (SU) network that has got to maximize the overall framework rate by selecting on ideal transmitting power value on each channel, and at the same moment, don't lead to any harmful interference to the receivers of the PU organize as investigates the radio resource allocation for a post- disaster surveillance system which is constructed using CRN through dynamic spectrum access (DSA) system using MAB.

E. FL-BASED SOLUTIONS

In [119], a multi-UAV system has been designed to study picture categorization in area exploration scenarios. FL has been used to complete image classification tasks, and local updates from all UAVs are broadcast to the ground fusion center (GFC) over fading wireless channels, based on the local model learned from pictures captured by an onboard camera at each UAV suggested FL-aided classification. As shown in Figure 2, FL is the least technique that is used in UAV problems, although it is a promising technique.

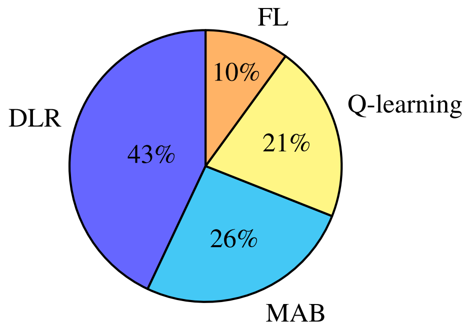


FIGURE 2. ML percentage usage in UAVs.

V. CHALLENGES AND FUTURE WORK

AI utilization for UAV systems has driven to present numerous development and savvy arrangements for an endless run of problems as shown in Figure 3. This section briefly surveys the major vital open subjects specified already for UAV issues summarized in Figure 4. From our study, it is clear that more than 40% of researchers used DRL because of its common policy and there are a lot of data to be utilized, but we believe that meta and federated learning will give better accurate and faster results if researchers develop it as it is a hot area to go through and find more methods to solve several problems. Drones are used to obtain confidential data, such as weather forecasting, storm tracking, and precision agriculture. They can even be used for surveillance purposes, especially in search and rescue. Future promising related research issues are as follow:-

- **UAV mounted RIS**: Recently, UAV-mounted RIS is under investigation to further improve wireless coverage and accuracy position. It is a promising research direction to leverage different AI techniques, especially online learning to enhance different related problems such as joint optimization and path planning [127].
- **Multi UAV path planning**: Most current UAV path planning handles a single UAV scenario with a static environment. Still, multi UAV trajectory planning within a dynamic environment via online learning is a vital future direction. For example, how to avoid obstacles, timing to select the best path, and prevent trajectory interference using AI optimization schemes.
- **UAV for V2X**: Although previous work utilized UAVs to make smart traffic control, the techniques used are DL-based, which consumes offline training time. It is promising to leverage self-learning or meta learning schemes proper for such difficulty.
- **Meta learning aided UAVs**: Since the number of users served by UAVs is increasing rapidly, we need to accelerate the learning process (i.e., reducing learning time) using advanced ML such as meta learning. This can

TABLE 6. MAB based UAV solutions.

Reference	Scenario Motivation	ML Scheme	Main Contribution	Limitations
[114] UAV-assisted emergency communications	BS in disaster Optimal path starting and ending	ϵ-greedy, and UCB algorithms	Provide emergency communication service for a post-disaster area with unknown user distribution	Limited battery capacity
[58] mmWave UAV wireless networks	Gateway selection maximize average data rate and minimizing battery cost	UCB, and TS algorithms	Maximize the achievable data rates of the access-gateway	Data loss
[115] Disaster surveillance system	Resource management (RM)	UCB algorithm	Secondary user network has to maximize the total system rate by selecting a proper transmitting power value on each channel, and on the same time, do not cause any harmful interference to the receivers of the PU network	Ideal setting
[116] 5G Beam selection for UAV applications	3D Beam selection	TS, ϵ-greedy, and Bayesian algorithms	Suitable beam selection technique	Non energy aware
[117] V2V communications	V2V channel selection in autonomous vehicle networks	UCB algorithm	Coordinate the operations of power control and channel selection while maintaining a desirable throughput	Simple network; inaccurate distance estimation

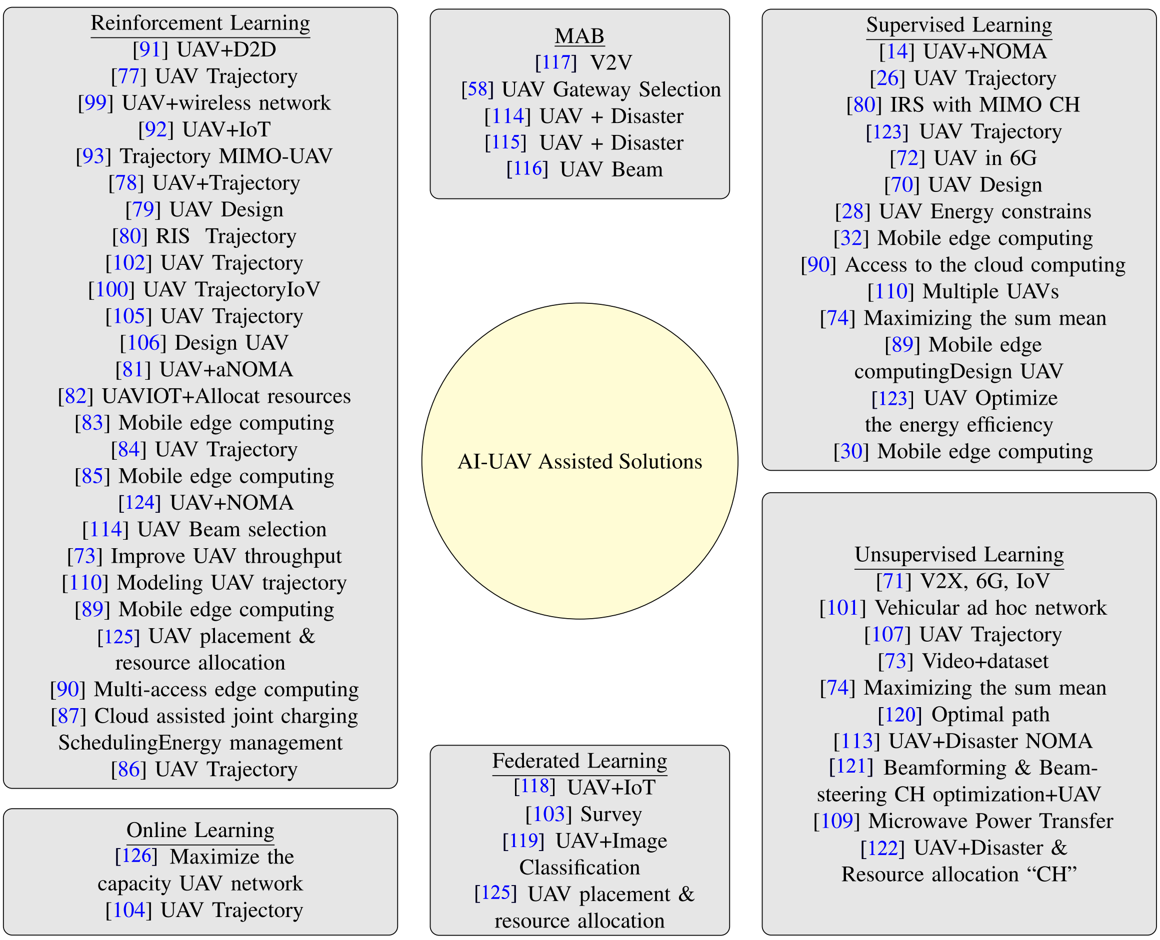


FIGURE 3. AI-assisted UAV applications.

enhance/accelerate node search methods, collaborative UAVs, and Cognitive Radio aided UAVs, and multipath planning.

- **UAV aided Wireless Power Transfer:** Recently, UAVs can be used to provide WPT to mobile devices that lacks energy. Hence, in that case online load balancing should

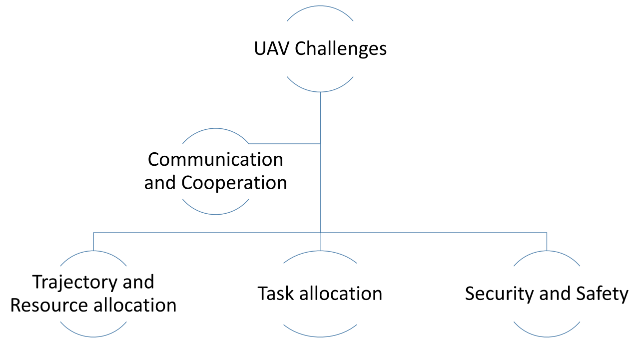


FIGURE 4. UAV challenges.

be implemented to distribute the energy fairly between the users. Such a topic is a promising future direction too, especially when optimizing the path of UAV to serve more users.

- **MEC aided UAVs**: Lately, both MEC and UAV are combined to simultaneously extend and facilitate UAV usage in different fields. Moreover, advanced AI techniques can enhance such a framework, especially FL ones at multi UAV dynamic scenarios with effective resource management.
- **Security/Privacy**: Designing the position of a UAV takes into account the existence of several eavesdroppers to enhance the secrecy performance. Enhancing and upgrading the current designs of cognitive anti-jamming V2V communications, and improving the inter-system relationship between radar tracking and vehicular communication. Advanced AI schemes should strengthen data security and user privacy against eavesdroppers.
- **Energy consumption**: Current related research aims to perform multi objective optimization to prolong the UAV battery lifetime to serve more users and maximize the sum rate. Hence, energy consumption is a open AI related issue that needs more investigation, especially for multi UAv scenarios.
- **UAV aided network caching**: Caching and computing can be incorporated into UAV-based integrated system to provide uRLLC in the emergencies. This can be done via up to date ML techniques proper to each specific scenario.

Ultimately, we suggest using online learning instead of offline learning in all previous research topics as it will be more efficient and with high response and accurate decisions on urgent issues.

VI. CONCLUSION

In this survey, we deeply investigated new ML-based research direction to improve the performance of UAV networks beneficial to a large variety of potential applications such as smart cities and airborne BS deployment, etc. Beforehand, we highlighted different ML types such as SL, USL, RL, FL, etc. Then, we surveyed distinct ML-aided UAV solutions according to the utilized ML category. Finally, we focused on MAB-assisted solutions as a promising direction due to various MAB types, proper to different scenarios. We offered a series of concluding observations for each of the strategies we looked at, outlining the existing limits and concerns as well as a set of interesting open problems. Finally, we summarized future directions and provided attractive UAV-related research topics that need more investigation, especially AI-aided ones.

REFERENCES

[1] X. Liu, Y. Liu, and Y. Chen, "Reinforcement learning in multiple-UAV networks: Deployment and movement design," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8036–8049, Aug. 2019.

[2] T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and N. Nasser, "UV-CDS: An energy-efficient scheduling of UAVs for premises sterilization," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 3, pp. 1191–1201, Sep. 2021.

[3] T. Tazrin, Z. Ayaz, M. M. Fouda, and Z. M. Fadlullah, "On energy-efficient UAV route scheduling to offload health data from under-served rural communities," in *Proc. IEEE Int. Conf. Internet Things Intell. Syst. (IoTaIS)*, Nov. 2021, pp. 86–91.

[4] M.-A. Lahmeri, M. A. Kishk, and M.-S. Alouini, "Stochastic geometry-based analysis of airborne base stations with laser-powered UAVs," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 173–177, Jan. 2020.

[5] O. M. Bushnaq, M. A. Kishk, A. Celik, M.-S. Alouini, and T. Y. Al-Naffouri, "Optimal deployment of tethered drones for maximum cellular coverage in user clusters," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 2092–2108, Mar. 2021.

[6] Y. Qin, M. A. Kishk, and M.-S. Alouini, "Performance evaluation of UAV-enabled cellular networks with battery-limited drones," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2664–2668, Dec. 2020.

[7] M. A. Kishk, A. Bader, and M.-S. Alouini, "On the 3-D placement of airborne base stations using tethered UAVs," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5202–5215, Aug. 2020.

[8] M. Kishk, A. Bader, and M.-S. Alouini, "Aerial base station deployment in 6G cellular networks using tethered drones: The mobility and endurance tradeoff," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 103–111, Dec. 2020.

[9] B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV assistance paradigm: State-of-the-art in applications and challenges," *J. Netw. Comput. Appl.*, vol. 166, Sep. 2020, Art. no. 102706.

[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[11] J. Konecný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016, *arXiv:1610.05492*.

[12] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 3rd Quart., 2019.

[13] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.

[14] P. S. Bithas, E. T. Michailidis, N. Nomikos, D. Vouyioukas, and A. G. Kanatas, "A survey on machine-learning techniques for UAV-based communications," *Sensors*, vol. 19, no. 23, p. 5170, 2019.

[15] A. Carrio, C. Sampedro, A. Rodriguez-Ramos, and P. Campoy, "A review of deep learning methods and applications for unmanned aerial vehicles," *J. Sensors*, vol. 2017, pp. 1–13, Aug. 2017.

[16] P. V. Klaine, R. D. Souza, L. Zhang, and M. Imran, "An overview of machine learning applied in wireless UAV networks," in *Wiley 5G Ref: The Essential 5G Reference Online*, 2019, pp. 1–15.

[17] M.-A. Lahmeri, M. A. Kishk, and M.-S. Alouini, "Artificial intelligence for UAV-enabled wireless networks: A survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1015–1040, 2021.

[18] S. Rezwan and W. Choi, "Artificial intelligence approaches for UAV navigation: Recent advances and future challenges," *IEEE Access*, vol. 10, pp. 26320–26339, 2022.

[19] R. Azoulay, Y. Haddad, and S. Reches, "Machine learning methods for UAV flocks management—A survey," *IEEE Access*, vol. 9, pp. 139146–139175, 2021.

[20] S. Zhang, H. Zhang, B. Di, and L. Song, "Cellular UAV-to-X communications: Design and optimization for multi-UAV networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1346–1359, Jan. 2019.

[21] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "UAV-enabled communication using NOMA," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5126–5138, Jul. 2019.

[22] J. Park, H. Lee, S. Eom, and I. Lee, "UAV-aided wireless powered communication networks: Trajectory optimization and resource allocation for minimum throughput maximization," *IEEE Access*, vol. 7, pp. 134978–134991, 2019.

[23] N. Zhao, X. Pang, Z. Li, Y. Chen, F. Li, Z. Ding, and M.-S. Alouini, "Joint trajectory and precoding optimization for UAV-assisted NOMA networks," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3723–3735, May 2019.

[24] W. Sun, G. Tang, and K. Hauser, "Fast UAV trajectory optimization using bilevel optimization with analytical gradients," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2020, pp. 82–87.

[25] W. Wang, J. Tang, N. Zhao, X. Liu, X. Y. Zhang, Y. Chen, and Yi Qian, "Joint precoding optimization for secure SWIPT in UAV-aided NOMA networks," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5028–5040, Aug. 2020.

[26] Z. Wang, F. Zhou, Y. Wang, and Q. Wu, "Joint 3D trajectory and resource optimization for a UAV relay-assisted cognitive radio network," *China Commun.*, vol. 18, no. 6, pp. 184–200, Jun. 2021.

[27] H. Jiang, R. He, C. Ruan, J. Zhou, and D. Chang, "Three-dimensional geometry-based stochastic channel modeling for intelligent reflecting surface-assisted UAV MIMO communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2727–2731, Dec. 2021.

[28] C. Su, F. Ye, L.-C. Wang, L. Wang, Y. Tian, and Z. Han, "UAV-assisted wireless charging for energy-constrained IoT devices using dynamic matching," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4789–4800, Jun. 2020.

[29] W. Chang, Z.-T. Meng, K.-C. Liu, and L.-C. Wang, "Energy-efficient sleep strategy for the UBS-assisted small-cell network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 5178–5183, May 2021.

[30] J. Zhang, L. Zhou, F. Zhou, B.-C. Seet, H. Zhang, Z. Cai, and J. Wei, "Computation-efficient offloading and trajectory scheduling for multi-UAV assisted mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2114–2125, Feb. 2020.

[31] Q. Wang, A. Gao, and Y. Hu, "Joint power and QoE optimization scheme for multi-UAV assisted offloading in mobile computing," *IEEE Access*, vol. 9, pp. 21206–21217, 2021.

[32] R. Chen, L. Cui, M. Wang, Y. Zhang, K. Yao, Y. Yang, and C. Yao, "Joint computation offloading, channel access and scheduling optimization in UAV swarms: A game-theoretic learning approach," *IEEE Open J. Comput. Soc.*, vol. 2, pp. 308–320, 2021.

[33] R. He and Z. Ding, *Applications of Machine Learning in Wireless Communications*. London, U.K.: IET, 2019.

[34] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[35] M. E. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine learning for 5G/B5G mobile and wireless communications: Potential, limitations, and future directions," *IEEE Access*, vol. 7, pp. 137184–137206, 2019.

[36] A. Alkhateeb, I. Beltagy, and S. Alex, "Machine learning for reliable mmWave systems: Blockage prediction and proactive handoff," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2018, pp. 1055–1059.

[37] S.-E. Chiu, N. Ronquillo, and T. Javidi, "Active learning and CSI acquisition for mmWave initial alignment," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2474–2489, Nov. 2019.

[38] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Multi-user hybrid beamforming relying on learning-aided link-adaptation for mmWave systems," *IEEE Access*, vol. 7, pp. 23197–23209, 2019.

[39] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, Antwerp, Belgium, May 2020, pp. 1–5.

[40] J. Cui, Z. Ding, and P. Fan, "The application of machine learning in mmWave-NOMA systems," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–6.

[41] V. Raghavan, L. Akhoondzadeh-Asl, V. Podshivalov, J. Hulten, M. A. Tassoudji, O. H. Koymen, A. Sampath, and J. Li, "Statistical blockage modeling and robustness of beamforming in millimeter-wave systems," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3010–3024, Jul. 2019.

[42] M. Elsayed, K. Shimotakahara, and M. Erol-Kantarci, "Machine learning-based inter-beam inter-cell interference mitigation in mmWave," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[43] Q. Zhao, "Multi-armed bandits: Theory and applications to online learning in networks," *Synth. Lectures Commun. Netw.*, vol. 12, no. 1, pp. 1–165, Nov. 2019.

[44] S. Hashima, B. M. ElHalawany, K. Hatano, K. Wu, and E. M. Mohamed, "Leveraging machine-learning for D2D communications in 5G/beyond 5G networks," *Electronics*, vol. 10, no. 2, p. 169, Jan. 2021.

[45] S. Hashima, K. Hatano, E. Takimoto, and E. Mahmoud Mohamed, "Neighbor discovery and selection in millimeter wave D2D networks using stochastic MAB," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1840–1844, Aug. 2020.

[46] S. Hashima, K. Hatano, H. Kasban, and E. M. Mohamed, "Wi-Fi assisted contextual multi-armed bandit for neighbor discovery and selection in millimeter wave device to device communications," *Sensors*, vol. 21, no. 8, p. 2835, Apr. 2021.

[47] S. Hashima, K. Hatano, E. Takimoto, and E. M. Mohamed, "Minimax optimal stochastic strategy (MOSS) for neighbor discovery and selection in millimeter wave D2D networks," in *Proc. 23rd Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Oct. 2020, pp. 1–6.

[48] S. Hashima, E. M. Mohamed, K. Hatano, and E. Takimoto, "WiGig wireless sensor selection using sophisticated multi armed bandit schemes," in *Proc. 13th Int. Conf. Mobile Comput. Ubiquitous Netw. (ICMU)*, Nov. 2021, pp. 1–6.

[49] E. M. Mohamed, S. Hashima, K. Hatano, S. A. Aldossari, M. Zareei, and M. Rihan, "Two-hop relay probing in WiGig device-to-device networks using sleeping contextual bandits," *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1581–1585, Jul. 2021.

[50] E. M. Mohamed, S. Hashima, K. Hatano, and S. A. Aldossari, "Two-stage multiarmed bandit for reconfigurable intelligent surface aided millimeter wave communications," *Sensors*, vol. 22, no. 6, p. 2179, Mar. 2022.

[51] S. Hashima, M. M. Fouda, Z. M. Fadlullah, E. M. Mohamed, and K. Hatano, "Improved UCB-based energy-efficient channel selection in hybrid-band wireless communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[52] S. Hashima, M. M. Fouda, S. Sakib, Z. M. Fadlullah, K. Hatano, E. M. Mohamed, and X. Shen, "Energy-aware hybrid RF-VLC multi-band selection in D2D communication: A stochastic multi-armed bandit approach," *IEEE Internet Things J.*, early access, Mar. 24, 2022, doi: 10.1109/JIOT.2022.3162135.

[53] M. M. Fouda, S. Hashima, S. Sakib, Z. M. Fadlullah, K. Hatano, and X. Shen, "Optimal channel selection in hybrid RF/VLC networks: A multi-armed bandit approach," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6853–6858, Jun. 2022.

[54] E. M. Mohamed, S. Hashima, K. Hatano, H. Kasban, and M. Rihan, "Millimeter-wave concurrent beamforming: A multi-player multi-armed bandit approach," *Comput., Mater. Continua*, vol. 65, no. 3, pp. 1987–2007, 2020.

[55] B. M. ElHalawany, S. Hashima, K. Hatano, K. Wu, and E. M. Mohamed, "Leveraging machine learning for millimeter wave beamforming in beyond 5G networks," *IEEE Syst. J.*, vol. 16, no. 2, pp. 1739–1750, Jun. 2022.

[56] S. Hashima, K. Hatano, H. Kasban, M. Rihan, and E. M. Mohamed, "Multiagent multi-armed bandit techniques for millimeter wave concurrent beamforming," in *Proc. 8th Int. Jpn.-Afr. Conf. Electron., Commun., Comput. (JAC-ECC)*, Dec. 2020, pp. 56–59.

[57] S. Hashima, Z. M. Fadlullah, M. M. Fouda, E. M. Mohamed, K. Hatano, B. M. ElHalawany, and M. Guizani, "On softwarization of intelligence in 6G networks for ultra-fast optimal policy selection: Challenges and opportunities," *IEEE Netw.*, early access, Feb. 18, 2022, doi: 10.1109/MNET.103.2100587.

[58] E. M. Mohamed, S. Hashima, A. Aldosary, K. Hatano, and M. A. Abdelghany, "Gateway selection in millimeter wave UAV wireless networks using multi-player multi-armed bandit," *Sensors*, vol. 20, no. 14, p. 3947, Jul. 2020.

[59] D. Li, S. Wang, H. Zhao, and X. Wang, "Context-and-social-aware online beam selection for mmWave vehicular communications," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8603–8615, May 2021.

[60] P. Zhou, J. Xu, W. Wang, C. Jiang, K. Wang, and J. Hu, "Human-behavior and QoE-aware dynamic channel allocation for 5G networks: A latent contextual bandit learning approach," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 2, pp. 436–451, Jun. 2020.

[61] S. Ali, A. Ferdowsi, W. Saad, N. Rajatheva, and J. Haapola, "Sleeping multi-armed bandit learning for fast uplink grant allocation in machine type communications," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5072–5086, Aug. 2020.

[62] Y. Xing, Y. Qian, and L. Dong, "A multi-armed bandit approach to wireless information and power transfer," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 886–889, Apr. 2020.

[63] C. Wang, J. Yang, H. He, R. Zhou, S. Chen, and X. Jiang, "Neighbor cell list optimization in handover management using cascading bandits algorithm," *IEEE Access*, vol. 8, pp. 134137–134150, 2020.

[64] P. Kairouz *et al.*, "Advances and open problems in federated learning," 2019, *arXiv:1912.04977*.

[65] Q.-V. Pham, M. Le, T. Huynh-The, Z. Han, and W.-J. Hwang, "Energy-efficient federated learning over UAV-enabled wireless powered communications," *IEEE Trans. Veh. Technol.*, vol. 71, no. 5, pp. 4977–4990, May 2022.

[66] J. Vanschoren, "Meta-learning," in *Automated Machine Learning* (The Springer Series on Challenges in Machine Learning), F. Hutter, L. Kotthoff, and J. Vanschoren, Eds. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-05318-5_2.

[67] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2020.

[68] A. Soltani and A. Izquierdo, "Adaptive learning under expected and unexpected uncertainty," *Nature Rev. Neurosci.*, vol. 20, no. 10, pp. 635–644, Oct. 2019.

[69] Q.-V. Pham, T. Huynh-The, M. Alazab, J. Zhao, and W.-J. Hwang, "Sum-rate maximization for UAV-assisted visible light communications using NOMA: Swarm intelligence meets machine learning," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10375–10387, Oct. 2020.

[70] H. E. Hammouti, M. Benjillali, B. Shihada, and M.-S. Alouini, "Learn-as-you-fly: A distributed algorithm for joint 3D placement and user association in multi-UAVs networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5831–5844, Dec. 2019.

[71] J. Hu, C. Chen, L. Cai, M. R. Khosravi, Q. Pei, and S. Wan, "UAV-assisted vehicular edge computing for the 6G Internet of Vehicles: Architecture, intelligence, and challenges," *IEEE Commun. Standards Mag.*, vol. 5, no. 2, pp. 12–18, Jun. 2021.

[72] R. Gupta, A. Shukla, and S. Tanwar, "BATS: A blockchain and AI-empowered drone-assisted telesurgery system towards 6G," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 4, pp. 2958–2967, Oct. 2021.

[73] S. Girisha, M. M. M. Pai, U. Verma, and R. M. Pai, "Performance analysis of semantic segmentation algorithms for finely annotated new UAV aerial video dataset (ManipalUAVid)," *IEEE Access*, vol. 7, pp. 136239–136253, 2019.

[74] Y. Wang, J. Yang, M. Liu, and G. Gui, "LightAMC: Lightweight automatic modulation classification via deep learning and compressive sensing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3491–3495, Mar. 2020.

[75] W. P. Amorim, E. C. Tetila, H. Pistori, and J. P. Papa, "Semi-supervised learning with convolutional neural networks for UAV images automatic recognition," *Comput. Electron. Agricult.*, vol. 164, Sep. 2019, Art. no. 104932.

[76] G. Afifi and Y. Gadallah, "Autonomous 3-D UAV localization using cellular networks: Deep supervised learning versus reinforcement learning approaches," *IEEE Access*, vol. 9, pp. 155234–155248, 2021.

[77] R. Ding, F. Gao, and X. S. Shen, "3D UAV trajectory design and frequency band allocation for energy-efficient and fair communication: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7796–7809, Aug. 2020.

[78] S. Zhu, L. Gui, N. Cheng, F. Sun, and Q. Zhang, "Joint design of access point selection and path planning for UAV-assisted cellular networks," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 220–233, Jan. 2020.

[79] W. Li, L. Wang, and A. Fei, "Minimizing packet expiration loss with path planning in UAV-assisted data sensing," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1520–1523, Dec. 2019.

[80] X. Guo, Y. Chen, and Y. Wang, "Learning-based robust and secure transmission for reconfigurable intelligent surface aided millimeter wave UAV communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1795–1799, Aug. 2021.

[81] T. Zhang, Z. Wang, Y. Liu, W. Xu, and A. Nallanathan, "Joint resource, deployment, and caching optimization for AR applications in dynamic UAV NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3409–3422, May 2022.

[82] A. M. Seid, G. O. Boateng, S. Anokye, T. Kwantwi, G. Sun, and G. Liu, "Collaborative computation offloading and resource allocation in multi-UAV-assisted IoT networks: A deep reinforcement learning approach," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12203–12218, Aug. 2021.

[83] L. Wang, K. Wang, C. Pan, W. Xu, N. Aslam, and L. Hanzo, "Multi-agent deep reinforcement learning-based trajectory planning for multi-UAV assisted mobile edge computing," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 1, pp. 73–84, Mar. 2021.

[84] A. Ferdowsi, M. A. Abd-Elmagid, W. Saad, and H. S. Dhillon, "Neural combinatorial deep reinforcement learning for age-optimal joint trajectory and scheduling design in UAV-assisted networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1250–1265, May 2021.

[85] L. Zhang, Z.-Y. Zhang, L. Min, C. Tang, H.-Y. Zhang, Y.-H. Wang, and P. Cai, "Task offloading and trajectory control for UAV-assisted mobile edge computing using deep reinforcement learning," *IEEE Access*, vol. 9, pp. 53708–53719, 2021.

[86] M. Samir, C. Assi, S. Sharafeddine, D. Ebrahimi, and A. Ghrayeb, "Age of information aware trajectory planning of UAVs in intelligent transportation systems: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12382–12395, Nov. 2020.

[87] S. Jung, W. J. Yun, M. Shin, J. Kim, and J.-H. Kim, "Orchestrated scheduling and multi-agent deep reinforcement learning for cloud-assisted multi-UAV charging systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5362–5377, Jun. 2021.

[88] H. Peng and X. Shen, "Multi-agent reinforcement learning based resource management in MEC- and UAV-assisted vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 131–141, Jan. 2021.

[89] H. Wang, H. Ke, and W. Sun, "Unmanned-aerial-vehicle-assisted computation offloading for mobile edge computing based on deep reinforcement learning," *IEEE Access*, vol. 8, pp. 180784–180798, 2020.

[90] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, and X. Shen, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.

[91] K. K. Nguyen, N. A. Vien, L. D. Nguyen, M.-T. Le, L. Hanzo, and T. Q. Duong, "Real-time energy harvesting aided scheduling in UAV-assisted D2D networks relying on deep reinforcement learning," *IEEE Access*, vol. 9, pp. 3638–3648, 2021.

[92] Z. Xiong, Y. Zhang, W. Y. B. Lim, J. Kang, D. Niyato, C. Leung, and C. Miao, "UAV-assisted wireless energy and data transfer with deep reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 1, pp. 85–99, Mar. 2021.

[93] R. Ding, Y. Xu, F. Gao, and X. Shen, "Trajectory design and access control for air–ground coordinated communications system with multiagent deep reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5785–5798, Apr. 2022.

[94] O. S. Oubbati, A. Lakas, and M. Guizani, "Multiagent deep reinforcement learning for wireless-powered UAV networks," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 16044–16059, Sep. 2022.

[95] O. S. Oubbati, M. Atiquzzaman, H. Lim, A. Rachedi, and A. Lakas, "Synchronizing UAV teams for timely data collection and energy transfer by deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6682–6697, Jun. 2022.

[96] S. Wu, W. Xu, F. Wang, G. Li, and M. Pan, "Distributed federated deep reinforcement learning based trajectory optimization for air-ground cooperative emergency networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 9107–9112, Aug. 2022.

[97] M. K. Shehzad, S. A. Hassan, A. Mahmood, and M. Gidlund, "On the association of small cell base stations with UAVs using unsupervised learning," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Apr. 2019, pp. 1–5.

[98] T. M. Hoang, N. M. Nguyen, and T. Q. Duong, "Detection of eavesdropping attack in UAV-aided wireless systems: Unsupervised learning with one-class SVM and K-means clustering," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 139–142, Feb. 2020.

[99] X. Guan, Y. Huang, C. Dong, and Q. Wu, "User association and power allocation for UAV-assisted networks: A distributed reinforcement learning approach," *China Commun.*, vol. 17, no. 12, pp. 110–122, Dec. 2020.

[100] A. Al-Hilo, M. Samir, C. Assi, S. Sharafeddine, and D. Ebrahimi, "UAV-assisted content delivery in intelligent transportation systems-joint trajectory planning and cache management," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5155–5167, Aug. 2021.

[101] S. Jiang, Z. Huang, and Y. Ji, "Adaptive UAV-assisted geographic routing with Q-learning in VANET," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1358–1362, Apr. 2021.

[102] H. Wu, F. Lyu, C. Zhou, J. Chen, L. Wang, and X. Shen, "Optimal UAV caching and trajectory in aerial-assisted vehicular networks: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2783–2797, Dec. 2020.

[103] B. Brik, A. Ksentini, and M. Bouaziz, "Federated learning for UAVs-enabled wireless networks: Use cases, challenges, and open problems," *IEEE Access*, vol. 8, pp. 53841–53849, 2020.

[104] S. Chai and V. K. N. Lau, "Online trajectory and radio resource optimization of cache-enabled UAV wireless networks with content and energy recharging," *IEEE Trans. Signal Process.*, vol. 68, pp. 1286–1299, 2020.

[105] O. Bouhamed, H. Ghazzai, H. Besbes, and Y. Massoud, "A UAV-assisted data collection for wireless sensor networks: Autonomous navigation and scheduling," *IEEE Access*, vol. 8, pp. 110446–110460, 2020.

[106] S. Zhu, L. Gui, D. Zhao, N. Cheng, Q. Zhang, and X. Lang, "Learning-based computation offloading approaches in UAVs-assisted edge computing," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 928–944, Jan. 2021.

[107] X. Liu, Y. Liu, Y. Chen, and L. Hanzo, "Trajectory design and power control for multi-UAV assisted wireless networks: A machine learning approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7957–7969, Aug. 2019.

[108] T. Zhang, Z. Wang, Y. Liu, W. Xu, and A. Nallanathan, "Caching placement and resource allocation for cache-enabling UAV NOMA networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12897–12911, Nov. 2020.

[109] K. Li, W. Ni, E. Tovar, and A. Jamalipour, "On-board deep Q-network for UAV-assisted online power transfer and data collection," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 12215–12226, Dec. 2019.

[110] J. Cui, Z. Ding, Y. Deng, A. Nallanathan, and L. Hanzo, "Adaptive UAV-trajectory optimization under quality of service constraints: A model-free solution," *IEEE Access*, vol. 8, pp. 112253–112265, 2020.

[111] T. Yang, Z. Jiang, R. Sun, N. Cheng, and H. Feng, "Maritime search and rescue based on group mobile computing for unmanned aerial vehicles and unmanned surface vehicles," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7700–7708, Dec. 2020.

[112] W. Shafik, S. M. Matinkhah, S. S. Afolabi, and M. N. Sanda, "A 3-dimensional fast machine learning algorithm for mobile unmanned aerial vehicle base stations," *Int. J. Adv. Appl. Sci.*, vol. 10, no. 1, pp. 28–38, 2020.

[113] W. Feng, J. Tang, N. Zhao, Y. Fu, X. Zhang, K. Cumanan, and K.-K. Wong, "NOMA-based UAV-aided networks for emergency communications," *China Commun.*, vol. 17, no. 11, pp. 54–66, Nov. 2020.

[114] Y. Lin, T. Wang, and S. Wang, "UAV-assisted emergency communications: An extended multi-armed bandit perspective," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 938–941, May 2019.

[115] A. Amrallah, E. M. Mohamed, G. K. Tran, and K. Sakaguchi, "Radio resource management aided multi-armed bandits for disaster surveillance system," *IEICE Proc. Ser.*, vol. 63, pp. 1–4, Dec. 2020.

[116] W. Shafik, M. Ghasemzadeh, and S. M. Matinkhah, "A fast machine learning for 5G beam selection for unmanned aerial vehicle applications," *J. Inf. Syst. Telecommun.*, vol. 4, no. 28, p. 262, 2020.

[117] S. Feng and S. Haykin, "Cognitive risk control for anti-jamming V2V communications in autonomous vehicle networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9920–9934, Oct. 2019.

[118] W. Y. B. Lim, S. Garg, Z. Xiong, Y. Zhang, D. Niyato, C. Leung, and C. Miao, "UAV-assisted communication efficient federated learning in the era of the artificial intelligence of things," *IEEE Netw.*, vol. 35, no. 5, pp. 188–195, Sep./Oct. 2021.

[119] H. Zhang and L. Hanzo, "Federated learning assisted multi-UAV networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14104–14109, Nov. 2020.

[120] K. Peng, W. Liu, Q. Sun, X. Ma, M. Hu, D. Wang, and J. Liu, "Wide-area vehicle-drone cooperative sensing: Opportunities and approaches," *IEEE Access*, vol. 7, pp. 1818–1828, 2018.

[121] L. Li, H. Ren, Q. Cheng, K. Xue, W. Chen, M. Debbah, and Z. Han, "Millimeter-wave networking in the sky: A machine learning and mean field game approach for joint beamforming and beam-steering," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6393–6408, Oct. 2020.

[122] B. Wang, Y. Sun, N. Zhao, and G. Gui, "Learn to coloring: Fast response to perturbation in UAV-assisted disaster relief networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3505–3509, Mar. 2020.

[123] M. Aloqaily, O. Bouachir, A. Boukerche, and I. A. Ridhawi, "Design guidelines for blockchain-assisted 5G-UAV networks," *IEEE Netw.*, vol. 35, no. 1, pp. 64–71, Jan./Feb. 2021.

[124] S. K. Mahmud, Y. Liu, Y. Chen, and K. K. Chai, "Adaptive reinforcement learning framework for NOMA-UAV networks," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 2943–2947, Sep. 2021.

[125] Q. V. Do, Q.-V. Pham, and W.-J. Hwang, "Deep reinforcement learning for energy-efficient federated learning in UAV-enabled wireless powered networks," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 99–103, Jan. 2022.

[126] X. Zhong, Y. Guo, N. Li, Y. Chen, and S. Li, "Deployment optimization of UAV relay for malfunctioning base station: Model-free approaches," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 11971–11984, Dec. 2019.

[127] E. M. Mohamed, S. Hashima, and K. Hatano, "Energy aware multiarmed bandit for millimeter wave-based UAV mounted RIS networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 6, pp. 1293–1297, Jun. 2022.

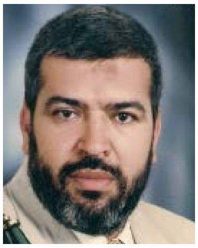
AMIRA O. HASHESH received the B.Sc. and M.Sc. degrees in electrical engineering from the Faculty of Engineering at Shoubra, Benha University, Egypt, in 2013 and 2020, respectively, where she is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Faculty of Engineering at Shoubra. Her research interests include wireless communications, machine learning, online learning, 5G, and the Internet of Things.

SHERIEF HASHIMA (Senior Member, IEEE) received the B.Sc. degree (Hons.) in electronics and communication engineering (ECE) from Tanta University, Egypt, in 2004, the M.Sc. degree (Hons.) in electronics and communication engineering (ECE) from Menoufiya University, Egypt, in 2010, and the Ph.D. degree from the Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt, in 2014. He has been a Postdoctoral Researcher with the Computational Learning Theory Team, RIKEN-AIP, Japan, since July 2019. He has been working as an Assistant then Associate Professor at the Engineering and Scientific Equipment Department, Nuclear Research Center (NRC), Egyptian Atomic Energy Authority (EAEA), Egypt, since 2014. From January 2018 to June 2018, he was a Visiting Researcher at the Center for Japan-Egypt Cooperation in Science and Technology, Kyushu University. He is a technical committee member in many international conferences and a reviewer in many international conferences, journals, and IEEE TRANSACTIONS. His research interests include wireless communications, machine learning, online learning, 5G, B5G, 6G systems, image processing, millimeter waves, nuclear instrumentation, and the Internet of Things. He is a member of AAAI.

ROKAIA M. ZAKI received the master's degree and the Ph.D. degree in communications and electronics engineering from Benha University, Egypt, in 2007 and 2012, respectively. She is currently an Assistant Professor with the Faculty of Engineering at Shoubra, Benha University. Her research interests include mobile communications, wireless networks, and cognitive radio.

MOSTAFA M. FOUDA (Senior Member, IEEE) received the Ph.D. degree in information sciences from Tohoku University, Japan, in 2011. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Idaho State University, ID, USA. He also holds the position of a Full Professor with Benha University, Egypt. He has worked as an Assistant Professor with Tohoku University. He was a Postdoctoral Research Associate with Tennessee Technological University, TN, USA. He has more than 110 publications in international conferences, journal articles, and book chapters. His research interests include cyber security, machine learning, blockchain, the IoT, 6G networks, smart healthcare, and smart grid communications. He has served on the technical committees of several IEEE conferences. He is also a Reviewer in several IEEE Transactions and Magazines. He is an Editor of IEEE Transactions on Vehicular Technology (TVT) and an Associate Editor of IEEE Access.

KOHEI HATANO received the Ph.D. degree from the Tokyo Institute of Technology, in 2005. Currently, he is an Associate Professor at the Research and Development Division, Kyushu University Library. He is also the Leader of the Computational Learning Theory Team, RIKEN AIP. His research interests include machine learning, computational learning theory, and online learning and their applications.



ADLY S. TAG ELDIEN received the B.Sc., M.Sc., and Ph.D. degrees from Benha University, Egypt, in 1984, 1989, and 1993, respectively. He is currently the Ex-Head of the Network and Information Center and the Head of the Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University. His research interests include robotics, networks, and mobile communication.