

Received December 29, 2021, accepted February 10, 2022, date of publication March 8, 2022, date of current version March 16, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3157651

Machine Learning Approaches for Reconfigurable Intelligent Surfaces: A Survey

K. M. FAISAL, (Graduate Student Member, IEEE), AND WOYEOL CHOI[✉], (Member, IEEE)

Department of Computer Engineering, Chosun University, Gwangju 61452, Republic of Korea

Corresponding author: Wooyeol Choi (wyc@chosun.ac.kr)

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2021R111A3050535.

ABSTRACT Next-generation wireless networks must handle a growing density of mobile users while accommodating a rapid increase in mobile data traffic flow and a wide variety of services and applications. High-frequency waves will perform an essential role in future networks, but these signals are easily obstructed by objects and diminish over long distances. Reconfigurable intelligent surfaces (RISs) have attracted considerable interest because of their potential to improve wireless network capacity and coverage by intelligently changing the wireless propagation environment. Consequently, RISs possess potential technology for the sixth generation of communication networks. Machine learning (ML) is an effective method for maximizing the possible advantages of RIS-assisted communication systems, particularly when the computational complexity of operating and deploying RIS increases rapidly as the number of interactions between the user and the infrastructure starts to grow. Since ML is a promising strategy for improving a network and its performance, the application of ML in RISs is expected to open new avenues for interdisciplinary studies as well as practical applications. In this paper, we extensively investigate the ML algorithms used in RISs. We provide a brief overview of RISs, a summary of ML methods with RIS architecture, and a comparison of the available methodologies to explain the combination of these two technologies. Moreover, the significance of open research topics is emphasized to provide sound research directions.

INDEX TERMS Reconfigurable intelligent surface, machine learning, deep learning, federated learning, reinforcement learning.

I. INTRODUCTION

Web-enabled gadgets, such as smartphones, have emerged as vital tools for global communication, information transfer, and entertainment. The academia and industry are now focusing on sixth generation wireless technology as the wireless sector is in a highly exciting moment where the fifth generation (5G) technology has been largely standardized and commercialized. According to the Cisco Annual Internet Report (2018-2023), mobile connectivity is expected to be available to more than 70% of the global population by 2023 and the number of overall mobile subscribers is expected to increase from 5.1 billion in 2018 to 5.7 billion in 2023 [1]. Inter-cell synchronization approaches have been built to solve the interference as cellular networks have become denser owing

The associate editor coordinating the review of this manuscript and approving it for publication was Nikhil Padhi[✉].

to more aggressive frequency reuse. However, the bandwidth of a network is still constrained owing to the irregularity of wireless transmission and accessible spectrum [2].

The limited availability of spectrum for communication systems is encouraging a gradual migration towards the higher frequency bands with abundant unoccupied spectra. However, as the radio frequency increases, the electromagnetic (EM) waves become more susceptible to obstruction from objects such as buildings in metropolitan regions. Adding more relays and base stations (BSs) to minimize communication distances and provide better network coverage consumes more energy. As a result, employing traditional cellular methods to assure wireless service coverage is challenging. To deal with the spectrum scarcity of communication systems, reconfigurable intelligent surfaces (RISs) have evolved as an important wireless network resolution for attaining high spectrum and energy efficiency [3].

For upcoming wireless communication networks such as beyond 5G, the RIS is projected as viable technology with the potential to significantly increase link quality and minimize the possibility of blockages. Small, low-cost passive components are piled together in the RIS to reflect incoming signals with a controllable phase shift toward the receiver. The comparatively simple deployment of RIS-assisted communications with affordable passive parts makes them valuable in smart radio contexts.

However, some certain challenges must be addressed before obtaining the advantages of RISs. Accurate channel state information (CSI) for optimum reflection on the RIS is required. It is challenging for a realistic RIS-aided wireless network to achieve a precise value for CSI on a continuous basis because of the potential flexibility of the served client and the obstruction-prone character of the signal. Consequently, the issues of CSI assessment and optimization of network performance under poor CSI should be appropriately addressed to allow real-time and effective RIS-assisted transmission. Owing to the utilization of considerable number of components, channel assessment complexity is high in RIS-assisted wireless networks, which is a major challenge; moreover, obtaining channel knowledge may require a large training overhead. Furthermore, the phase shift of the reflecting elements complicates the designing of an ideal passive beamforming system, and the conventional methodologies require complicated procedures for the configuration of the RIS which is both power and time consuming.

Owing to their ability to learn and the requirement of operating over wider search areas, machine learning (ML) techniques have attracted attention in wireless communications [4]–[8], especially in the field of RISs. Over the last few years, several researchers have attempted to overcome these obstacles. They have been working with various ML algorithms for the communication sector so that the infrastructure can independently solve all challenges. Most ML methods work by learning the parameters and constructing an optimization model from the input information for the goal function. In the present arena, as a massive amount of data must be handled, the efficiency and effectiveness of mathematical optimization procedures significantly impact the popularity and application of ML models [9].

Although few studies have been performed on the application of artificial intelligence (AI) in RISs, in the literature, there is no survey that exists exclusively based on ML applications in RISs. To overcome this gap, our survey offers a comprehensive assessment of the state-of-the-art applications of ML in RISs. Subsequently, the techniques are classified according to the optimization targets. A comparative study was conducted among all the reviewed techniques.

As the first step of this study, in [10], we presented a brief introduction to RIS and a simplified introduction to the machine learning techniques used in RIS. However, this paper contains a general and detailed description of RIS, and various machine learning-based algorithms such as supervised learning, unsupervised learning, reinforcement learning, and

federated learning applied to RIS systems are explained in depth. Furthermore, a more detailed comparison with the advantage and drawbacks of each technique is provided in this paper. The contributions of our study are as follows.

- A concise overview of the RIS architecture is provided for an important insight into this evolving architecture.
- A brief introduction to ML is presented, and different ML techniques are introduced.
- Existing surveys and studies on RISs and ML are presented for a better understanding.
- Applications of the ML techniques that have been used in the reviewed studies are revealed with the investigation, and the examined schemes are categorized based on their optimization goals and models.
- Finally, research issues are emphasized to offer valuable directions for future research along with the key challenges in employing ML in RISs.

The remainder of this paper is organized as follows. In Section II, an overview of the structural design of the RIS is discussed. Section III introduces ML designs that have been applied in the literature. Related works are addressed in Section IV. The applications of ML in RISs are described in Section V. The key research challenges and related future research issues are discussed in Section VI. Finally, in Section VII, this article is concluded with an outline of the entire work.

II. OVERVIEW OF RECONFIGURABLE INTELLIGENT SURFACE

RIS models are primarily created using metamaterials, which are periodically aligned subwavelength elements capable of providing complete control over EM actions of the metasurface and consist of unit cells [11]–[13]. This man-made EM material surface can be controlled electrically via integrated electronics and has unique wireless communication characteristics [14]. More precisely, an RIS functions by the placement of a large number of low-cost antenna components with the goal of controlling re-radiation and capturing energy. In the literature, varactor-centered and positive-intrinsic-negative diode based control methods were the standard techniques used [15]–[17]. To enhance the user communication quality and improve the properties of incident waves, control signals are transmitted by a BS to an RIS controller in an RIS-supported wireless network. The RIS does not perform digitizing because it operates as a reflector. Consequently, if properly implemented, the energy consumption of the RIS will be significantly lower than that of standard relays such as amplify-and-forward relay [18]–[20]. As illustrated in Figure 1, the practical EM wave-based tasks that RISs can employ in wireless communications are as follows:

- **Reflection:** An impacting radio wave is reflected in a particular direction, which may not be in the same direction as the incidence wave direction.
- **Refraction:** An impacting radio wave is refracted which may not be in the same direction as the incidence wave.

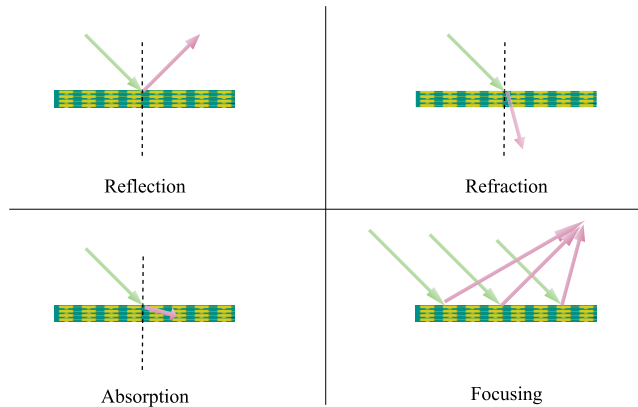


FIGURE 1. Electromagnetic wave-based activities of a reconfigurable intelligent surface.

- **Absorption:** This entails creating a smart surface that cancels the refracted and reflected radio waves corresponding to a certain incident radio wave.
- **Focusing:** It entails directing an impinging radio beam to a certain point.

A. PERSPECTIVE OF PHYSICS

EM waves encounter dispersed particles while traveling across space, which attenuates the signal. The physics-based bedrock of surface electromagnetism is the surface equivalence theorem. The Huygens principle asserts that each point across a wavefront is a generator of spherical wavelets, and additional wavelets emerging from various sites overlap. The wavefront is formed by the addition of several spherical wavelets. The EM field radiated by an RIS can be computed and analyzed based on the Huygens principle.

Figure 2(a) [21], [22] illustrates a volume V occupied by several EM radiation sources consisting of charges q_i and currents J_i . Just outside the volume V , these sources generate a magnetic induction field B and an electric field E . The arrangement of scatterers can be substituted by an arbitrarily thin layer of particular magnetic currents J_m and electric currents J_e that completely covers the volume V , as per the Huygens principle. Magnetic currents can only be created by cycles of electric currents with a limited depth. Hence, the layer thickness can be electrically negligible but not zero. EM fields are scattered exclusively outside the volume V by the corresponding surface currents J_m and J_e , and all these EM fields are identical to those formed by the original sources. Huygens’ surfaces that are related to currents that disperse EM fields solely with one side may be extended to metamaterials.

The boundary conditions are based on the fact that when an average tangential field is applied to a thin sheet of polarizable objects, it induces magnetic J_{ms} and electric J_{es} surface currents, which may be linked to the applied fields using magnetic surface admittance Y_m and electric surface impedance Z_e . Figure 2(b) [23]–[25] shows a magnetic surface admittance $Y_m(x,y)$ and an electric surface impedance $Z_e(x,y)$, which define the physical configuration

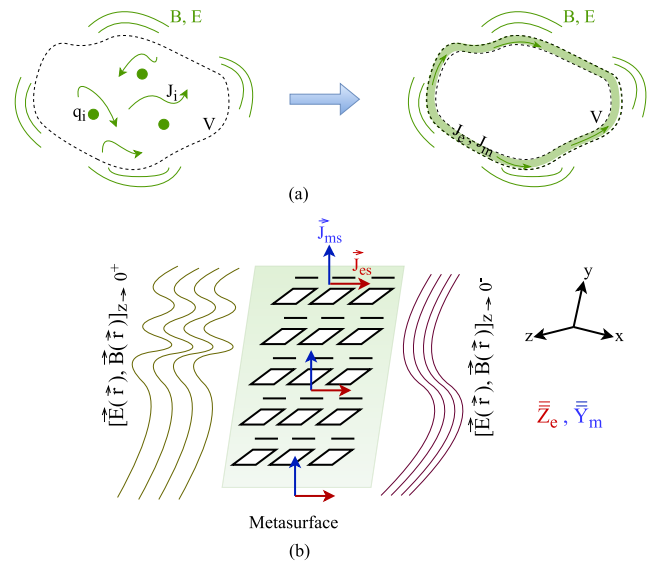


FIGURE 2. Representation of (a) surface equivalent theorem scattered EM source and (b) physical layout of the metasurface.

of a generic sheet of the metasurface. The mean applied field induces magnetic and electric currents on the metasurface, creating a discontinuity between the fields above and below the surface, thereby allowing wavefront modification.

B. INTERACTION BETWEEN THE CELLS

The RIS modulation is dependent on the intercell connection of tunable chips, which regulate the scattering components of the metasurface to provide the desired tunable functions. Wireless or cable communication is possible among the underlying chip controllers. Because wired communication is easier to combine with the controllers on the same chip, it is a better option; however, in a significantly compact or large sized metasurface, wireless intercell communication is an effective solution. With strict robustness requirements and energy latency, the design guidelines for inter-communication procedures must be practiced [26]. The exact application is determined by either the size of the tile or the desired wavelength. Two separate connection pathways are shown in the Figures 3 (a), (b) [27]. In case (a), the metasurface layer, which is the gap between the plane at the back and the metasurface patches, is the first channel. The antenna is a part of the chip, whereas the role of the waveguide is performed by the plane at the back and metasurface patches. In case (b), a separate control plane is constructed by inserting additional metal slabs beneath the chip for the second channel. As in the aligned-plate waveguide, monopoles supplied from the chip could generate waves that travel in this barrier condition.

C. RELATIONSHIP BETWEEN THE METASURFACE AND THE RIS

Metasurface is a two-dimensional planar metamaterial with EM properties. Metamaterials have not been discovered in natural supplies, and they are composed several tightly placed

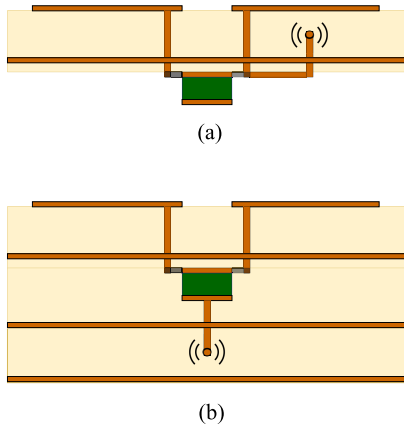


FIGURE 3. Communication channels: (a) metasurface substrate, (b) waveguide with specific parallel plates.

subwavelength resonating structures known as meta-atoms or pixels [28]. The distinguishing characteristics are their ability to shape EM waves in a variety of ways. Owing to their petite size, a significant number of these closely packed atoms provide large degrees of freedom in altering the incident EM waves. For instance, a metasurface can impose arbitrary quasi-continuous [29] amplitude or phase profiles on the incident wavefronts and exert fine-grained control over the dispersed electric field by carefully incorporating its meta-atoms.

In general, software and meta-atom oriented controllers are essential elements of the RIS that influence the metasurface reconfiguration rate. The related power consumption of static and reconfigurable metasurfaces is significantly different because no active electrical circuits are required for static metasurfaces; they can be completely passive. As energy is required to control the received signals and switches for reconfiguration, metasurfaces with reconfigurable properties can only be virtually passive. However, a specialized power supply is not required for signal transmission after the metasurface has been appropriately calibrated.

D. PASSIVE BEAMFORMING AND RIS

When multiple antennas produce identical signal copies of the postponed signal, beamforming occurs. Constructive interference occurs in geographic places where the signal copies are collected simultaneously, whereas at other places, destructive interference occurs. When multiple antenna send signals, the receiver will collect better signals than when a single antenna transmits signals while consuming the same total power. The time delays at the transmitting antennas are set to create constructive interference at the receiver. This traditional array gain demonstrates that the beamformed signal becomes more spatially concentrated if there is an increase in array size. The received signal strength and surface area are proportional, and depend on the number of elements of the transmitter. With the delay in time, when the RIS re-radiates the chosen signal, an array gain is produced to beamform the signal at the receiver, similar to the traditional manner.

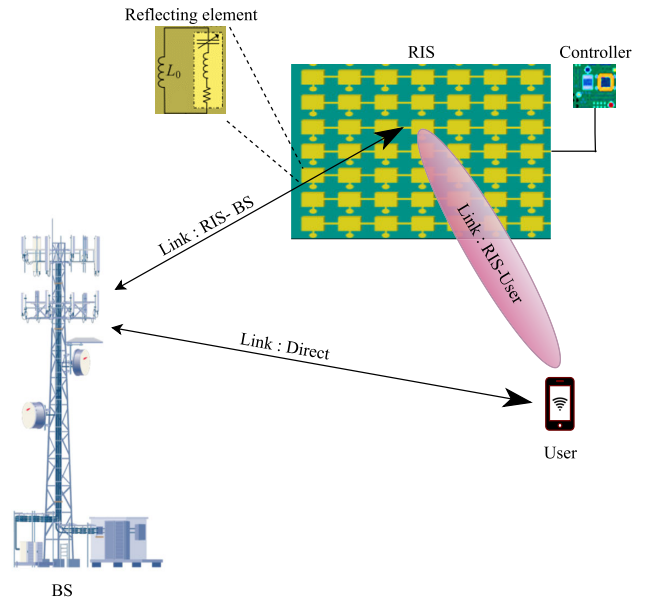


FIGURE 4. Passive beamforming using an RIS.

The process of passive beamforming by the RIS between the BS and the user by reflecting the signals to aid in communication is shown in Figure 4. The RIS reflection coefficients can be modified by the BS using an RIS controller. Furthermore, passive beamforming at the RIS and transmit beamforming at the BS must be developed together to increase communication performance [30].

III. FUNDAMENTALS OF MACHINE LEARNING

A section of science that studies the theory and characteristics of learning algorithms, their performance, and associated systems is known as ML. ML is a wide multidisciplinary area that draws concepts from a variety of domains, including information theory, AI, statistics, optimal control, optimization theory, and a variety of other scientific, mathematical, and engineering disciplines [31]–[34]. ML has touched nearly every scientific subject owing to its deployment in diverse applications, which has a significant influence on research and society [35]. Currently, ML is predominantly applied in autonomous systems, suggestion engines, informatics, data mining, and recognition systems [36]. The ML technique typically comprises two major phases: training and decision making. In the training phase, a dataset is used to train and understand the model of the system. During the decision-making process, the trained model is employed to derive the projected output for every new input given to the system. The classification of ML involves various subfields such as reinforcement learning (RL), unsupervised learning, and supervised learning [37], as shown in Figure 5.

A. LEARNING ALGORITHM

1) SUPERVISED LEARNING

A type of learning that recognizes the parameters in the presence of a supervisor is called supervised learning.

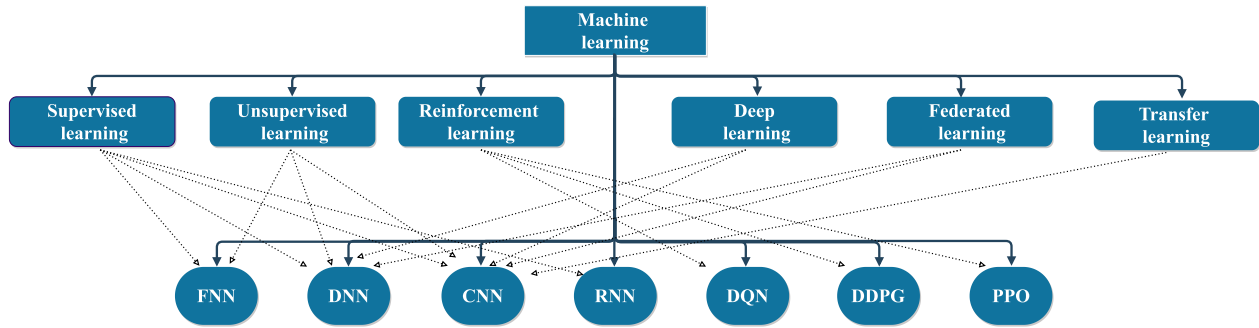


FIGURE 5. Classification of machine learning.

A collection of data is provided to an algorithm that includes both output and input information in this form of learning. A model for the data may be developed based on the output-input connection; then, to make a prediction, a fresh data set is input into the model [38], [39].

2) UNSUPERVISED LEARNING

A type of ML where, without the assistance of a supervisor, an algorithm provides the error level or right solutions for every inspection, with the aim of correctly understanding the outputs and series of inputs. In summary, an unsupervised learning technique obtains an unlabeled input dataset and effectively discovers the data connections to form a cluster [39].

3) REINFORCEMENT LEARNING

This is a commonly applied and effective ML method that learns about the environment by performing various actions and determining the best operation strategy. The two fundamental factors of RL are the environment and the agent. By applying the Markov decision process (MDP) [40], the agent investigates the surroundings and determines the action that must be implemented for the optimum result.

Q-learning (QL) [41] is a straightforward and effective RL method in which a model of the environment is not required; the goal is achieved based on the reward. The process of updating the Q-values for an RL task can be expressed as follows:

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \phi \max_a Q(s+1, a) - Q(s, a)]$$

When action a is chosen, $Q(s, a)$ is the current value of the state s ; $0 < \alpha < 1$ is the learning constant and $0 < \phi < 1$ is the discounting factor. The algorithm operates as follows: the agent chooses an action at some state s . Given that the action a is implemented, it discovers the highest feasible Q-value in the following state ($s+1$) and changes the current Q-value. The discounting factor provides the choice of either rewarding in the future ($\text{if } \phi \gg 0$) or presenting immediate rewards ($\text{if } \phi \ll 1$). To improve the convergence and stability of the algorithm, a constant is used to adjust the learning rate. QL has been previously used in various wireless situations,

such as in wireless sensor network routing [42]–[44]. It is simple to set up and exhibits an acceptable balance between memory and energy needs.

4) DEEP LEARNING

A subset of ML is deep learning (DL), which allows an algorithm to generate projections and classifications without being explicitly programmed based on the decisions of input data. Some cases of DL include QL, k-nearest neighbor classifiers, and linear regression. DL algorithms may extract information from raw data in a hierarchical manner by utilizing nonlinear processing components of multiple layers for forecasting outcomes based on the desired objective [45]. Recently, DL has attracted more interest from the academic community because of its superior performance in areas such as computer vision, information retrieval, speech recognition, and language processing [46]–[49]. As computing power and graphic processors are improving daily [50], it is increasingly becoming important for areas involving big data sets to deliver projected analytic solutions.

5) FEDERATED LEARNING

This system typically includes an aggregated server and a wide range of devices, as illustrated in Figure 6. Every device trains a local model while the aggregating server updates and manages the global model. Through a series of communication sessions, the global and local models are modified repeatedly until global acceptance is established. Only devices with proper channel scenarios and relevant model updates are selected for updating and training the local model by the aggregating server. A localized model is updated by every device based on the most current downloaded global model. Then, the local modifications are transmitted to the aggregation server through wireless uplink data transfer to update the global model. This global model is then transmitted to the chosen devices through downlink broadcasts for further learning.

6) TRANSFER LEARNING

This is a learning algorithm that can enhance the performance of a new task using information from previously

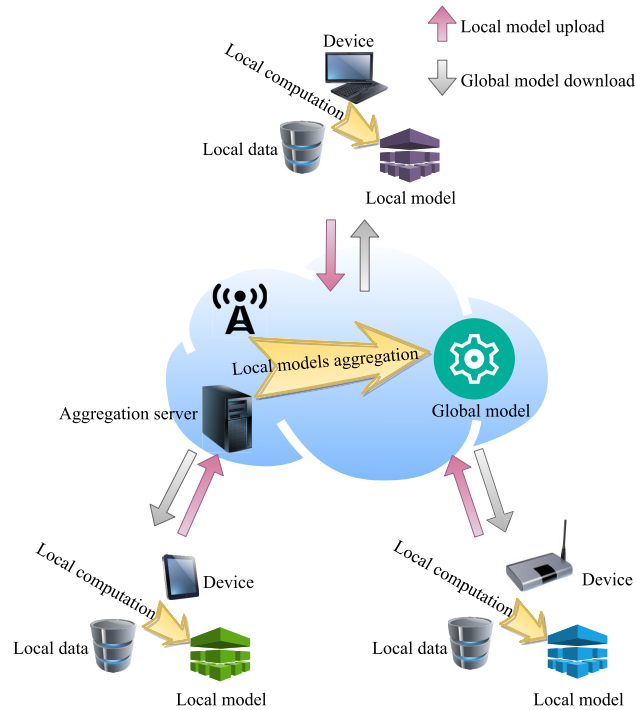


FIGURE 6. Federated learning scheme.

acquired tasks. Because many real-world problems do not require large amounts of labeled data sets to train sophisticated models, this algorithm is extremely beneficial in the field of data science. By utilizing data from the source, the objective of transfer learning (TL) is to enhance learning in a specified job [51].

B. NEURAL NETWORK

1) FEED-FORWARD NEURAL NETWORK

It consists of several hidden layers with a single layer each for input and output. Another name for the feed-forward neural network (FNN) model is multilayer perceptron (MLP), which is a simple DL model in which every neuron is linked with the neurons in the neighboring layer but not to the neurons of the same layer. The units of each layer are closely connected, necessitating the construction of a large number of weights. Figure 7 illustrates the architecture of the FNN, which shows that all neurons between two successive layers are completely linked. An effective algorithm for teaching an FNN with gradient descent is the backpropagation technique [52].

2) DEEP NEURAL NETWORK

The major objective of the deep neural network (DNN) is to learn from the data without performing manual calculations every time. A DNN is a nonlinear computational model that is structured similar to the human brain, which can learn and execute tasks such as decision-making, prediction, classification, and visualization tasks [53]. The architecture of the DNN [54] contains several layers of neurons, typically one linked output layer, more than one hidden layer, and one input

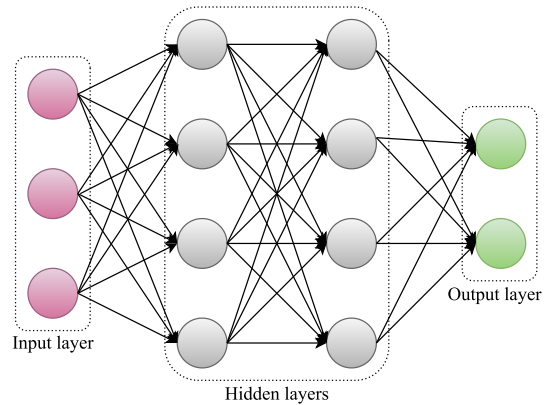


FIGURE 7. Feed-forward neural network.

layer. The input layer receives input from the input neurons and forwards it to the hidden layers. Through the hidden layer, the data are subsequently sent to the output layer. A weighted input, an activation function, and an output are present in every neuron. The output is determined by the activation function, which is dependent on the input of the neuron [55]. The distinction between a DNN and a neural network (NN) is that a shallow NN has one layer, whereas a DNN contains many hidden layers, where each layer of the DNN has a large number of neurons.

3) CONVOLUTIONAL NEURAL NETWORK

This network is built to handle data that are organized into numerous arrays [56]. The foundation of the convolutional NN (CNN) architecture is the use of convolutional layers to extract high-level features from two-dimensional data structures. The CNN is a class of the FNN [57], [58]. The three main principles of the CNN are pooling, weight sharing, and local sparse connections among the subsequent layers. These three basic principles substantially minimize the difficulty of training CNNs. Weight sharing refers to the fact that all neurons in the same convolution layer have the same weight parameters. Using weight sharing and local sparse connections, a reduction can be achieved. Pooling may be used to minimize the feature size while preserving the feature invariance.

4) RECURRENT NEURAL NETWORK

In recurrent NN (RNN), there are feedback connections between the connected neurons. Owing to their internal states, RNN can track the periodic connection of inputs and have one or more feedback connections between them. When a neuron in one layer sends data to the previous layer, it is called a feedback link [59]. Long short-term memory (LSTM), a popular form of the RNN [60], [61], can obtain long-term needs as it has a strong capacity. To calculate the hidden state, the LSTM employs three gates: a forget gate, an output gate, and an input gate. The LSTM can manage noise and dispersed representations of continuous data while

bridging a significantly long time lag. In LSTM, to control the cell condition and determine the outcome, the forget gate is utilized.

5) DEEP Q-NETWORK

The deep Q-network (DQN) is formed by the incorporation of a Q-table in a DNN. Similar to the deep reinforcement learning (DRL) technique, DQN is a QL technique that works with Q-values. DQN works to mitigate overestimation issues that frequently occur when agents move in complicated settings [62].

6) DEEP DETERMINISTIC POLICY GRADIENT

The deep deterministic policy gradient (DDPG) is a model-free actor-critic method. It is an off-policy method that combines the benefits of policy gradients and DQN [63]. Deterministic policies can be optimally learned in a high-dimensional continuous action space. Outputs and inputs are received by the actor-network as continuous action and then fed into the critic-network along with the state space. To determine an action, the actor network is at work, removing the requirement for non-convex optimization to identify the action that maximizes the subsequent state of the Q-value function.

7) PROXIMAL POLICY OPTIMIZATION

For situations with continuous or discrete action spaces, a policy gradient approach known as proximal policy optimization (PPO) can be used. It deploys the actor-critic approach and an on-policy approach to train a stochastic policy. The critic predicts reward of the agent by observing the actor from the knowledge obtained through the performed action. A series of courses was first gathered for every epoch by sampling the latest edition of a stochastic strategy. As a final step, the policy was updated by computing the reward system and the estimated benefits [64].

IV. EXISTING WORKS

In the literature, a few brief magazine papers, surveys, and tutorials introduce the RIS and its modifications; however, the focus of these works differs from ours. Several different terms and acronyms are often used in the literature to refer to RISs. In this study, we will use the term RIS to address all types of intelligent surfaces that are used in different studies. Hence, this section contains a summary of the existing works related to RISs.

RIS-enabled wireless networks, including the possible use of RIS in multiple-input multiple-output (MIMO) transmitters to either obtain low complications or to sharpen radio frequency (RF) signals, and the distinctive properties of RISs, are described in [15]. Simple analytic models were also utilized, with an emphasis on the performance of in terms of the error and link budget. In [65], the authors discussed the primary uses of the RIS in the design of hardware, new signal models, and critical challenges of wireless communication. An overview of the current research on large intelligent

surface-aided wireless networks, fundamentals of radio containing reflected waves such as reflective relay, backscatter communication, and the basics of RIS technology was also presented [13]. ElMossallamy *et al.* discussed models of an appropriate channel for the applications of an accurate estimation ability [2]. These characteristics distinguish the optimization of RIS from the precoding designed for typical MIMO arrays by emphasizing potential possibilities and future problems.

Moreover, in [66], the authors thoroughly examined of the theoretical foundations of RIS and presented a current assessment of various performance indicators and analytical methodologies for characterizing the improvement in the performance of wireless networks aided by RIS. In [12], an extensive review of the primary technical enablers, the ongoing status of studies, emerging ideas of smart radio environments and their challenges in research, major operational rules, and intended possible implementations were presented; moreover, in [30], delivered an unbiased view on RIS technology by examining the principles and then explaining specific aspects that can be readily misunderstood, thereby debunking a few myths.

Another survey [29] offered an introduction of communications in holographic MIMO (HMIMO), focusing on the significant obstacles in forming HMIMO-enabled wireless communications, thus emphasizing the potential and the reconfiguration of such surfaces with accessible hardware designs. Further, another study [20] explained the fundamental similarities and distinctions of RISs that are configured to act like reflectors or relays along with the numerical findings that emphasize the spectrum efficiency of RISs when compared to the radio wavelength.

In [67], the authors discussed the relationship between RL and an RIS-enabled software-controlled environment, as well as the notion of Wireless 2.0, which proposes significant modifications to the existing wireless network. Moreover, the importance of DL-focused RIS technology in communication systems and certain future research directions that are aimed at providing diverse technical discussions are presented [68]–[70].

Di Renzo *et al.* [22] provided a detailed assessment of an idea related to smart propagation environments using RIS and the potential applications, open research problems, primary operational concepts, and research advancements in this field. Tang *et al.* [71], described RIS-aided multi-stream transmitter models and the benefits of the metasurface-based receiver and RF chain-free transmitter for a potential energy-efficient and cost-effective communication technology.

Furthermore, the principles along with the obstacles in obtaining reliable communication networks, new opportunities, and future research issues of RIS-assisted communication were discussed [72], [73]. Equally important are the problems in the primary physical layer when integrating wireless networks and RISs. Yuan *et al.* addressed the issues of passive information transmission, CSI, a low-complexity

resilient design process, upcoming investigation scope, and security schemes in [74].

V. APPLICATIONS OF ML IN RIS

In this section, research works based on ML applications in the RIS architecture are presented with a comparative analysis. As shown in Figure 8, the ML frameworks used in RIS, such as resource management, security, beamforming, channel estimation, and various other aspects are examined thoroughly along with the key concept. In Table 1, ML approaches are listed according to their optimization targets and models.

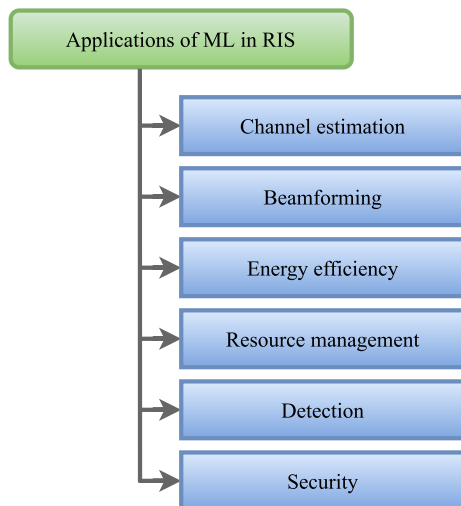


FIGURE 8. Various applications of ML in RIS.

A. CHANNEL ESTIMATION FOR RIS

In [96], the authors described a technique based on RL for maximizing throughput with both imperfect and perfect CSI. A quantile regression distributional reinforcement learning (QR-DRL) technique was used for each pair of action-state and to construct a return distribution, which approximated the innate unpredictability in the collaboration of the MDP between the environment and the RIS.

As illustrated in Figure 9, the environment is represented by the connected channels in the RL framework, and the agent is the controller of the RIS that executes action to adjust the coefficient of reflection. The RIS receives a communication reward after each time slot, which is characterized as the sum rate of the downlink in the direction of the user. Owing to the effect of fading in small-scale millimeter-wave (mmWave) channels, the CSI of the downlink may fluctuate even with a constant RIS coefficient of reflection. After observing a deviation vector at the conclusion of each time slot, the RIS strives to change its coefficient of reflection to fit the real downlink. To increase future rewards, a strategy is first developed to determine the chance by altering the coefficient of reflection in the current state to assess the possibility of every matrix in action. The capability of every action-state

pair in enhancing the transmission for the downlink is then calculated for each stationary policy. The distributional RL framework proposed by the authors differs from traditional RL tactics such as DQL [107]. In conventional RL methods, a scalar quantity is assigned to predict the return of the future, whereas in the aforementioned method proposed by these authors, the sum rate is modeled as a distribution function with the goal of considering for any doubt in the reward function.

Elbir *et al.* [78] developed a DL approach for the estimation of the channel in a massive MIMO system enhanced by RIS. The authors demonstrated that a method based on CNN achieves a stronger performance with a smaller normalized mean square error (NMSE) than the previous benchmark methods. The received signal is input to a double CNN with nine layers for the approximation of the cascaded and direct channels. The input layer acknowledges the received signals. The remaining layers are composed of a convolutional layer containing 256 filters with a size of 3×3 , and the final layer is the regression layer. A hyperparameter tuning procedure is subsequently incorporated to produce the optimum performance while keeping the network parameters fixed [108]–[110]. A supervised DL architecture was employed by plotting the received pilot signals with respect to the channels. Each user has the same CNN that receives the input pilot signals and estimates the explicit channel between the receiver and the transmitter. The proposed method can also be expanded to a multiple-user condition where each user has their own CNN and can assess their channel. Although retraining is unnecessary when there are changes in the position of the user, there is a requirement for additional control while manipulating the RIS components, and the training overhead is also high.

Liu *et al.* [77] presented a deep denoising NN-aided channel state prediction for RIS in mmWave systems to decrease the training overhead. Initially, the authors presented a hybrid RIS design, in which relatively limited RF receiver chains were used for achieving a trade off between hardware complications and performance. The authors presented a complex-valued denoising CNN (CV-DnCNN) to improve the estimation accuracy by exploiting the MIMO system with a delay in the angular domain channel matrix, which is produced by a denoising CNN (DnCNN) [19]. The suggested CV-DnCNN uses a network design similar to that of DnCNN [111], with the exception of complicated signal processing components, which may jointly process the imaginary and real parts of the channel matrix of an angular-delay domain, leveraging their relationship for improved performance. The authors incorporate the complicated structure blocks, which were inspired by [112], for complicated signal processing into the DnCNN for the modification of the denoiser. As shown in Figure 10, the components of the CV-DnCNN are an input layer, an output layer, and 15 convolutional layers. The primary layer consists of 64 filters, and for activation, a rectified linear unit (ReLU) is utilized. To accelerate the learning process

TABLE 1. Comparison of machine learning techniques for reconfigurable intelligent surfaces.

Algorithm	Ref.	Model	Layer information	Input	Output	System model	User	Goal
Supervised learning	[75]	CNN	2 convolutional, 2 fully connected	Incident RF signal	Set of interfering users	MISO	Multiple	Sum-rate maximization
	[76]	DNN	4 fully connected	Channel vector	Predictions on the achievable rates	OFDM-based single receiving antenna	Single	Achievable rate maximization
	[77]	CV-DnCNN	15 convolutional	Received pilot signals	Residual noise	mmWave mMIMO	Single	Performance maximization
	[78]	Twin CNN	3 convolutional, 2 fully connected	Received pilot signals	Channel matrices	mMIMO	Multiple	Performance maximization
	[79]	DeepRIS	3 fully connected	Symbols	Transmitted vectors	MISO	Single	Optimal BER
	[80]	DNN	3 fully Connected	User location	Optimal phase matrices	MISO downlink	Multiple	SNR maximization
	[81]	NN	1 hidden layer	RSSI	Mapping	SISO	Single	Higher precision
	[82]	ODE- CNN	26 layers	Sampled channel	Channel matrix	OFDM-based single receiving antenna	Single	Performance maximization
	[83]	CNN	41 layers	Sampled channel	Extrapolation of full channels	OFDM-based SISO	Single	Achievable rate maximization
	[83]	FNN	4 hidden layers	Sampled channel	Optimal beamforming vector	OFDM-based SISO	Single	Achievable rate maximization
	[84]	DL-MDC	3 hidden layers	Channel matrix	Beamforming matrix	Downlink THz mMIMO	Multiple	Sum Rate maximization
	[85]	RNN	2 layers of GRU	Beam and location sequences	Optimal beamforming vector	THz drone communication	Single	Optimal beam prediction
	[86]	DNN	3 hidden layers	Measured coordinate	Beam management	mmWave beam management	Multiple	Performance maximization
[87]	MLP	3 hidden layers	Sampled channel	Optimum phase	OFDM-based single receiving antenna	Single	Achievable rate maximization	
Unsupervised learning	[88]	LPSNet	1 or 2 hidden layers	Channel state	Phase shift	MIMO downlink	Single	Spectral efficiency maximization
	[89]	RISBFNN	5 fully-connected	Channel vector	Predicted phase	Single receiving antenna	Single	Gain enhancement
	[90]	CNN, FNN	2 convolutional, 3 fully-connected	Feature vector and effective channel	Phase prediction and beamforming matrix	MISO downlink	Multiple	Sum-rate maximization
Reinforcement learning	[91]	DDPG	2 hidden layers	Channel information	Q-function	MISO	Multiple	Sum-rate maximization
	[92]	DQN	4 layers	Channel vector	Interaction vector	OFDM-based single receiving antenna	Multiple	Achievable rate maximization
	[93]	DDPG	4 layers	State vector	Optimal phase shift matrices	MISO	Single	SNR maximization
	[94]	D3QN	3 fully connected	State vector	Beamforming policy	MISO downlink	Multiple	Energy efficiency maximization
	[95]	PDS-PER	3 layers	Channel vector	Beamforming policy	MISO downlink	Multiple	Secrecy rate enhancement
	[96]	QR-DRL	-	State vector	Optimum policy	MISO	Multiple	Sum rate maximization
	[97]	DDPG	Size of hidden layer is 48	Environment state	Resource allocation	MISO downlink NOMA	Multiple	Sum rate maximization
	[98]	DDPG	2 hidden layers	Environment state	Beamforming matrix, Q-function	Multi-hop THz	Multiple	Coverage range maximization
	[99]	DQN	2 fully-connected hidden layers	Channel values	Optimize RIS configuration	Uplink Multi-RIS	Multiple	Sum rate maximization
	[100]	DQN	3 linear layers	State vector	Resource scheduling	Vehicular communications	Multiple	Achievable rate maximization
	[101]	PPO	3 hidden layers	Activation patterns	Phase-shift matrix of aerial RIS	Aerial RIS	Multiple	Age-of-information sum minimization
	[102]	DDPG	-	State	Q-value	MISO	Single	Transmit power minimization
	[103]	D-DQN	3 layers	State	Q-function	MISO downlink	Multiple	Optimized beamforming and UAV trajectory
Federated learning	[72]	CNN	10 layers	Received pilot signals	Channel matrix	mMIMO	Multiple	UAV trajectory
	[104]	CNN	7 layers	Randomly generated data	Image classification	Multi-RIS	Multiple	Propagation error reduction
	[105]	OBR-FL	6 fully connected	Channel information	Beamformer weights	Single receiving antenna	Multiple	Data rate enhancement
Deep learning	[100]	CDRN	3 denoising block	Noisy channel matrix	Denosed channel matrix	TDD-based single receiving antenna	Multiple	Noise removal
	[106]	FNN	2 hidden layers	Channel vector	Quantization level	Single antenna user in TDD mode	Multiple	Enhancement of KGR in PHY key generation

and enhance denoising performance, batch normalization is utilized between the ReLU and the convolution layer. Unlike the traditional DnCNN, CV-DnCNN employs a complex convolutional layer.

The centralized learning (CL) method, in which the entire dataset is transmitted from the clients to the BS, has

a significant transmission overhead. Typically, after the model has been trained at the BS, the parameters obtained are delivered to the subscribers, who may then utilize the model to execute channel estimation tasks by providing it with the pilot data they have received. However, in [113], FL techniques for channel estimation across RIS-aided

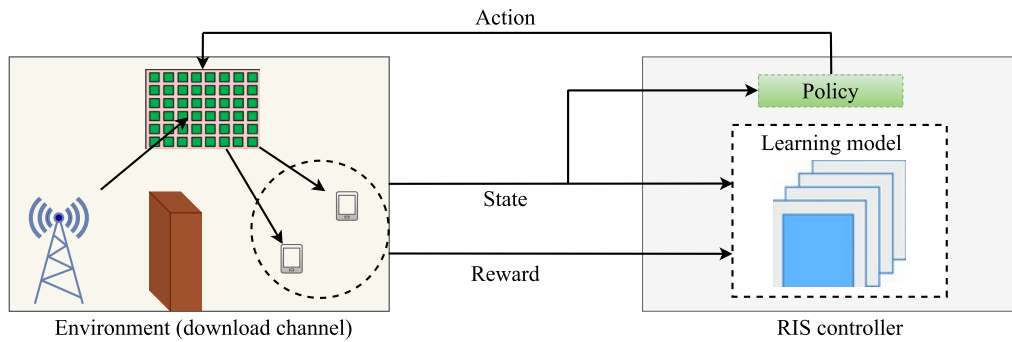


FIGURE 9. RIS-assisted downlink transmission based on a reinforcement learning framework.

massive MIMO are presented to address the significant transmission overhead of CL strategies. The author developed a CNN at the BS that was trained on local datasets. The collection of user data, training of a global model, and prediction of the own channel are the three stages of the proposed scheme. For both the RIS-aided massive MIMO and traditional operations, a single CNN (ChannelNet) is trained on two separate datasets, and a CNN with 10 layers is the suggested network design. Based on simulations, the suggested approach has shown a reduced transmission overhead when compared to the CL systems while retaining channel prediction performance that is comparable to CL, which exhibits only a minor estimation error. FL schemes are beneficial in minimizing significant transmission overhead; however, the reliability of the FL is typically worse than that of the CL when the model is trained only once.

Liu *et al.* [114] used a CNN to allow deep residual learning in RIS-aided systems to address the problem of restricted channel estimation performance and presented a CNN-based deep residual network (CDRN) for channel estimation. A CNN-based denoising block with a component-wise subtraction architecture is specifically intended to concurrently utilize both the cumulative characteristics of the noise and the spatial characteristics of the noisy channel matrices. The authors discussed a multi-user communications network with one RIS, several users, and one BS that uses the time division duplex (TDD) protocol. The CDRN is composed of denoising blocks, and a CNN with a deduction architecture is used in the denoising block to acquire the residual noise from the noisy channel matrix. Furthermore, as the channel matrix is complex, the input is split into imaginary and real parts for ease of feature extraction. The proposed technique achieved nearly the same prediction accuracy as that of the best minimal MSE (MMSE) estimator that relies on a prior probability density function of the channel, according to the simulation findings. However, the training overhead may be significant because of the large amount of data from multiple users.

Xu *et al.* [82] suggested an ordinary differential equation (ODE)-oriented CNN for estimating the cascaded channel in which the RIS components could be partially switched off by spatial sampling, thereby improving resource usage and

reducing the duration of the pilot phase in channel estimation. The authors analyzed an interior scenario in which a multiple-antenna access point (AP) uses RIS reflection to connect with a single-antenna device. The suggested extrapolation technique efficiently compressed the large-scale RIS channel throughout the physical environment, according to the simulation findings. Furthermore, the ODE-based design demonstrates the potential to increase the efficiency of traditional CNNs efficiency by accelerating convergence.

In [83], Zhang *et al.* proposed two approaches. The first one is a CNN-based channel extrapolation network with active antennas that can function together for extrapolating entire channels using the predicted partial channels in the channel extrapolation; the second approach is an FNN with an active antenna selection network to immediately map the predicted incomplete channels to the optimum beamforming vector using the beam searching technique. To determine the best positions for the active RIS components for both systems, the probabilistic sampling theory was used. The authors considered a single-antenna receiver and transmitter in an RIS-assisted communication system. The output dimension of the beam-seeking network is significantly lower than that of the CNN-based channel extrapolation network, which requires high training overhead; therefore, FNN is used to discover the best vector for beamforming. The beam searching technique is significantly more resilient than the channel extrapolation system with fewer active antennas, and the optimum antenna choice is better than the uniform antenna choice, according to the simulation findings. Future research may include an expansion to multiple- user conditions.

B. BEAMFORMING

Except for a few active elements attached to the baseband of the RIS controller, all RIS elements are passive. In [76], the authors employed a DL technique to learn the reflection matrices of the RIS by sampling the experience of the channel with no information on the geometry array. As only a few channel estimation techniques have been studied for intelligent surface communication systems that are focused on DL, the authors offer two techniques for designing RIS reflection matrices, where the DL-oriented approach has the

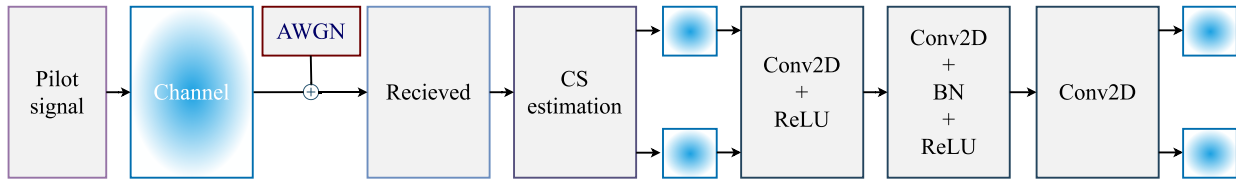


FIGURE 10. Model of complex-valued denoising convolutional neural network.

benefit of having minimal overhead, making it more viable. MLP networks are known as universal function approximators [115]; therefore, the application of a network based on MLP is encouraged to represent the relationship between the RIS interaction vector and the environment. The proposed method is thoroughly tested using a particular ray-tracing oriented dataset for DeepMIMO [116]. However, this study only examined labeled data to maximize the attainable single-user rate in their models and did not consider poor CSI, secure communication, and multiple user situations.

Taha *et al.* presented a DRL structure to eliminate labeled data with a low-overhead training that can adjust the phase shifts in an RIS [92]. The proposed method provides distributed RIS capability that may self-configure and operate with no aid from the BS or nodes in the infrastructure. The mathematical findings demonstrate that the recommended framework, when trained online, may achieve a performance similar to that of a perfect CSI. It is considered that the DQN receives the concatenated sampled channel vector as an input. For training stability, a double DQN is employed. The model is trained with the goal of forecasting the best vector of the interactions using a regression loss function. The proposed DRL-based algorithm seeks to maximize the possible rate of communication by the direct optimization of matrices in interaction based on the sampling knowledge of the channel. Each training episode in the recommended DRL structure uses only a single beam. Consequently, the training overhead can be eliminated, and the gathering period of the dataset is unnecessary.

To calculate the optimum RIS beamforming vector, a DL system is applied, as illustrated in Figure 11. The sampled vectors of the channel are called environment signifiers. These signifiers define the position of the receiver or transmitter and the nearby environment with some resolution. In DL, the algorithm attempt to discover a link between the observable environmental characteristics and the vectors of the best RIS reflection. This might be interpreted as identifying the interaction of RIS with the wireless signal based on the characteristics of the environment. The desired capability of an RIS is that vectors of a sampled channel should be produced with minimal learning costs. When compared to SL, even though RL allows for independent operation, it requires more training time.

In [80], Huang *et al.* suggested a supervised learning based technique for maximizing the power received in an RIS-aided network. In a multi-user, multiple-input single-output

(MISO) system, the downlink is studied, and a DNN is utilized to learn the plot between the formation associated with the RIS reflecting elements and the locations of users using a function approximator that maximizes the signal quality at every anticipated position of the user in an indoor situation. The proposed DNN architecture has five layers, where each layer of output possesses a nonlinear function.

As shown in Figure 12, on the floor of the room, the reference point and user location are assumed. An end-user is provided with a wireless device capable of conducting channel estimation by itself with the aid of the RIS and AP. It is anticipated that the propagation of the signal from the RIS and AP to the position of the desired user is achieved using a proper setup. Excluding the RIS, ray tracings are absorbed as they encounter the ceiling, floor, and walls of the interior environment. The dotted brown lines represent the absorbed signal beams. The proposed method in an indoor setting can lower the hardware burden while dealing with several BSs to enhance the signal; however, the method might not function well under various barriers in indoor circumstances.

Configuring reflecting factors without large channel estimation or the use of beam training is significantly difficult. In [87], the authors considered RIS-aided wireless networks and proposed a phase optimization technique by taking advantage of the correlation between the previously calculated and current channels. An MLP model was developed to enhance the quality of optimum RIS communication. Single-antenna transceivers communicate with each other. There are two steps in the proposed method. The first is a learning step in which the DL model is constructed and trained, and a testing stage in which the trained model is used to perform optimal phase interaction. For different conditions, the best RIS interaction performance was measured in terms of the attainable rate. Extensive experiments using a ray-tracing dataset demonstrated an increase in the attainable rate.

DL techniques were used by Gao *et al.* to minimize the model complications in RIS-aided networks in [89]. The authors offered an unsupervised label-free [117], [118] strategy for optimizing the phase shifts of RIS to minimize the labeling cost of supervised learning. By optimizing the goal function and preparing a DNN parameters, offline training with a modified DNN is proposed. The structural design of the NN comprises five entirely connected layers. Owing to the scaling of the system, the concept of proportionalizing the number of neurons to confirm the learning ability is considered. The initial four layers employ ReLU for the

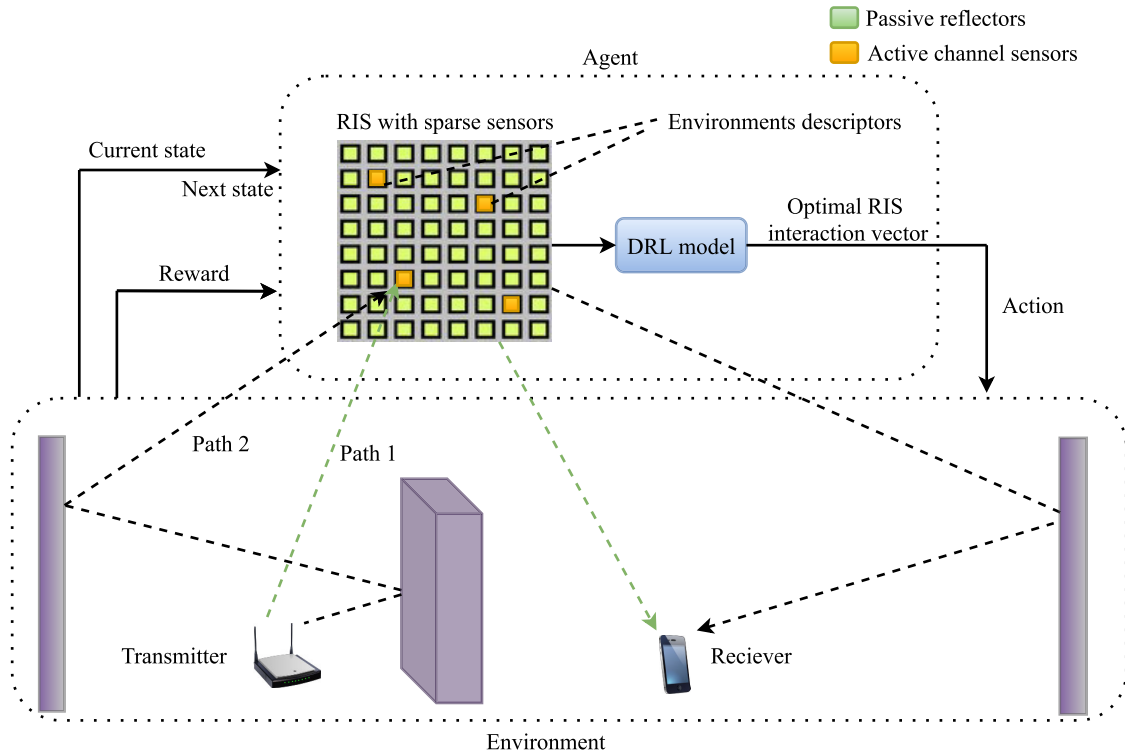


FIGURE 11. Idea of deep learning implementation in RIS.

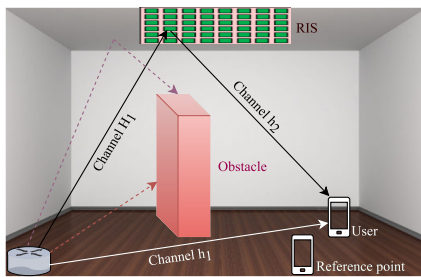


FIGURE 12. Indoor communication system with RIS.

activation function, whereas the last layer employs a linear unit to predict the phase shift. Even if the investigated technique remains empirical in the idea that no optimality of the property can be asserted, the simulation results demonstrate an improvement over standard methods based on the usage of alternating optimization (AO) and semi-definite relaxation.

An ideal joint optimization design is difficult because of the interaction between active and passive beamforming and the RIS phase changes. Song *et al.* presented a two-stage NN with an unsupervised learning-based solution for the joint passive and active beamforming design in RIS-assisted multi-user MISO downlink platforms to efficiently solve the joint optimization issue in [90]. Figure 13 illustrates the entire network design. For simplicity, the first and second halves of the network are referred to as PhaseNet and BeamNet. PhaseNet is used to anticipate RIS phase changes based on the

input characteristics, whereas BeamNet is used to forecast the beamforming matrix based on an effective communication channel built in the middle. Moreover, BeamNet serves as a PhaseNet analyzer. According to the simulation findings, the proposed method may reach an equivalent sum-rate capability with far less complexity than traditional optimization algorithms, enabling a real-time beamforming setup in RIS-assisted platforms.

For MIMO systems, a significant level of computational effort is required in the AO technique for the passive beamforming architecture and evaluation techniques in the RIS platform. Nguyen *et al.* [88] suggested an unsupervised learning technique with fewer complications involved in the RIS-assisted spectral efficiency optimization problem in the MIMO system. When generating the phase shifts for the RIS, the recommended scheme has a simplified input design and only demands a limited number of nodes and layers. The authors have considered a downlink transmission between a BS and a user-assisted by an RIS in this study, wherein the 8×2 MIMO system can forecast RIS phase shifts using only one hidden layer, but a 16×2 MIMO systems requires two hidden layers. In the simulation results, the spectrum efficiency obtained from the proposed method appears to be higher than of the AO technique.

In [119], Liaskos *et al.* proposed an NN framework for customizing the actions of tiles in RIS-aided systems. A tailored propagation of the signal is modeled using backpropagation. After a training period, the NN understands how to set the

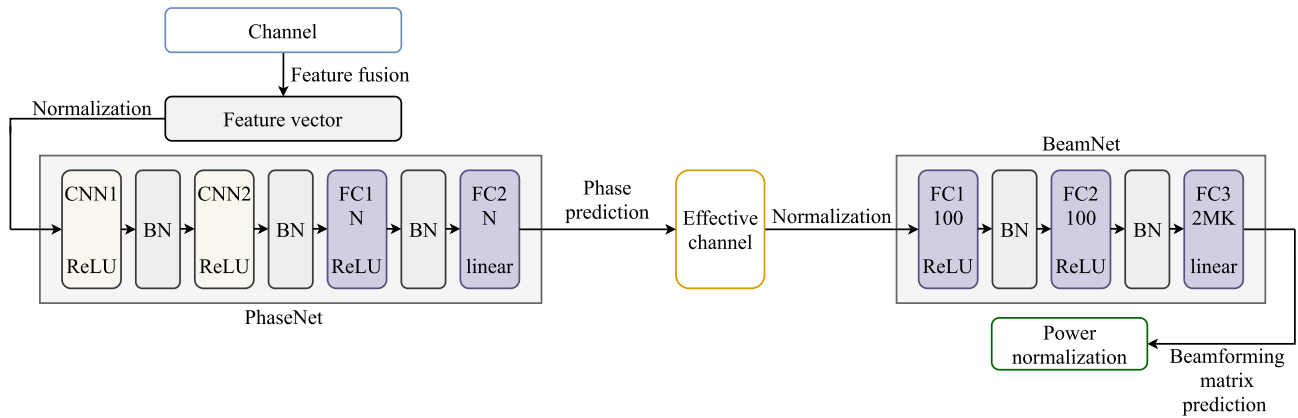


FIGURE 13. Architecture of a two stage network.

RIS tiles, providing improved performance. The propagation environments, which include the number and position of the receivers and transmitters, noise readings, operational frequencies, and the size of metasurface blocks, are among the data that are fed into the input layer. Backpropagation works by adjusting the node weights based on the errors at the output layer [120]. The fault can be split based on the intensity of connections between the nodes because each hidden node in the NN adds to the resulting error of its linked node to the output; thus, the resulting error is transmitted back to each layer for the weight modification, thereby minimizing the new error. The authors achieved consistency in performance gains by employing ray-tracing in a simulation environment.

Utilizing the structure of DRL, Huang *et al.* proposed a combined model of phase shifts in an RIS-aided system for MIMO and transmit beamforming [91]. The proposed DRL method for achieving scalability is proven to be useful for accommodating different system configurations. Instead of using the traditional AO method for acquiring RIS-aided phase shifts and transmit beamforming, the proposed methodology provides real-time improvement training of the DNN at the output of the model. The DDPG approach is incorporated by the authors using an immediate reward as a sum rate to maximize the throughput. The phase shift and constant transmit beamforming were simultaneously enhanced with minimal complications in the suggested model. The actor and critic networks are both completely linked DNNs with two hidden layers. The input and output dimensions of the actor network are specified as the cardinality of the state and action. In the components of the hidden layers, the neuron numbers are determined by the number of BS antennas, elements at the RIS, and users. In the recommended method, RIS can learn from the rewards and optimize itself according to the condition; however, owing to a lack of active elements, the availability of necessary data will be restricted, hampering its performance.

Blockage and channel limitations are common in THz telecommunications. To address these issues, Abuzainab *et al.* developed a THz drone system in which

a flying RIS and BS support a mobile drone user [85]. Figure 14(a), illustrates the THz technology scheme with RIS. It features a single-cell tower with a component array antenna that provides a beamforming vector to a mobile drone-based client. A preset beam codebook is used to determine the beamforming vectors. If the line of sight between the BS and the drone is disrupted, the drone is served by the BS via a hovering RIS. The service beams and record of the positions of the user are employed in a DL-based gated recurrent unit (GRU)-assisted RNN, as shown in Figure 14(b). The model calculates the optimal network connectivity (cascaded or direct) and optimum beamforming vector for that link ahead of time.

Owing to the vast number of reflecting components, passive beamforming may be constrained by high computational complexity. To overcome this issue, Gong *et al.* proposed an optimization-driven DRL strategy for collaborative beamforming, which is resistant to channel dynamics [121]. The RIS phase vector, power-splitting ratio, and the active beamforming vector of the AP are all parts of the joint beamforming optimization method for the RIS-aided MISO downlink system. The authors utilized an optimization-driven DDPG algorithm to build an optimization module to update and determine the best action in each decision epoch. As the DDPG algorithm accelerates the search process and reduces the search space, numerical findings show that reward performance and learning efficiency are considerably enhanced when compared to the traditional model-free DRL approach. However, supplemental optimization techniques add to the complexity of the procedure.

C. ENERGY EFFICIENCY FOR RIS

The combined optimization of the reflecting RIS and BS designs is difficult owing to the large number of RIS elements. Lee *et al.* presented a DRL-based solution to tackle this optimization problem, wherein the BS chooses the RIS ON/OFF states, allocation scheme, and phase shift [122]. Then, the environment modules, which include the user and RIS, transmit feedback information containing the status of

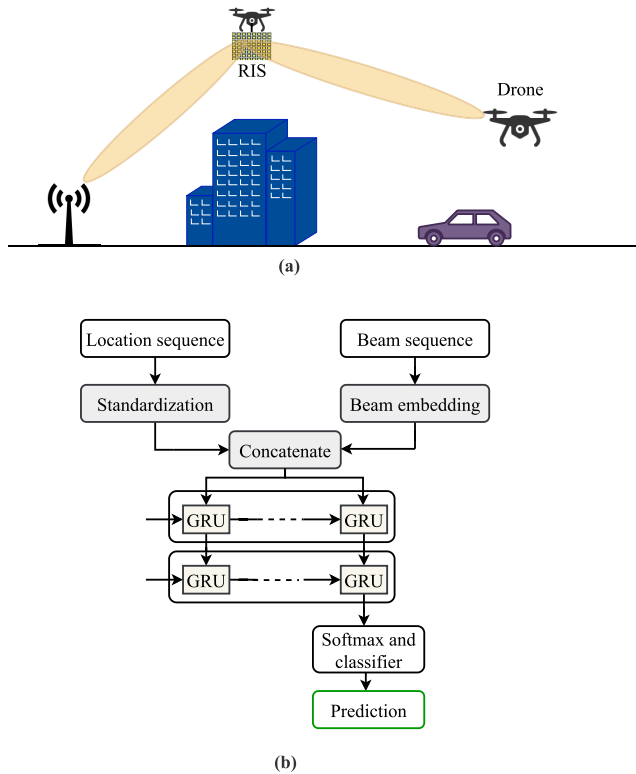


FIGURE 14. RIS in a THz communication scheme: (a) RIS with drone, (b) DL model to predict the operating RIS and BS.

the wireless channel and energy efficiency. During the suggested learning process, the BS can choose the best potential actions based on various states. The authors examined at a cellular network with a single BS downlink. As illustrated in Figure 15, the proposed architecture contains an agent installed on the BS and the remaining surrounding nodes, which include the users and RIS. The states are composed of the energy expenditure of the RIS and precoding vectors of the users. Moreover, optimization factors such as phase shifting, ON/OFF status, and transmit power are the actions. When the number of RIS components is increased, the simulation findings demonstrate that the proposed framework enhances energy efficiency.

In another study [84], the authors presented an RIS-aided hybrid precoding architecture for THz connectivity with reduced energy usage. To accomplish analog beamforming, the fundamental concept is to substitute the power-hungry array inside the standard hybrid precoding design with an energy-efficient RIS. To address the classification task with minor complications, a DL-based multiple discrete categorization (DL-MDC) hybrid precoding method is utilized. A conventional modern downlink THz massive MIMO system with 1-bit RIS phase shifts was studied, which can be readily built using reduced power and low-cost diodes. To cover a single-antenna client, the BS uses RF chains. Several DNNs may be used in a parallel DNN system in which every DNN has a single output that corresponds to

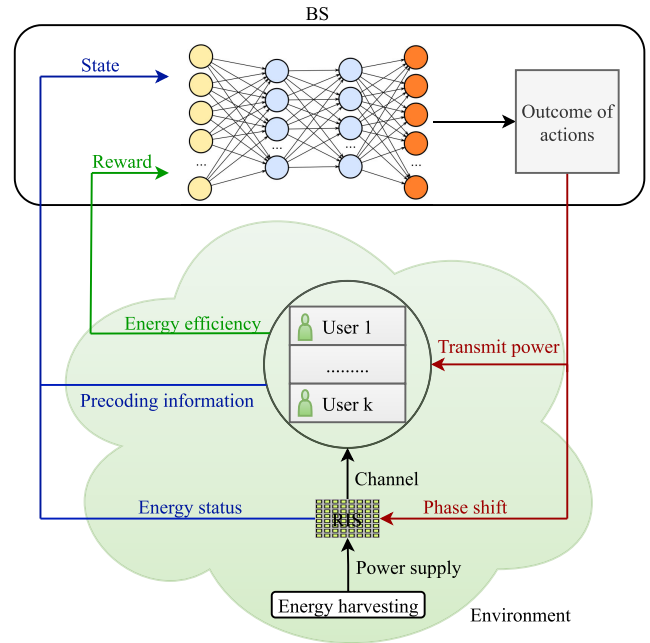


FIGURE 15. Energy efficiency model.

one diagonal member of the analog beamforming matrix. Consequently, all DNNs in the parallel DNN platform may be trained and employed simultaneously, thereby reducing the runtime significantly. Both the actual Third Generation Partnership Project channel model and the theoretical Saleh–Valenzuela channel model were used to develop the DL-MDC algorithm. It can reasonably mimic the complicated nonconvex function of the classic hybrid precoding algorithm, resulting in a good exchange between complexity and performance. In the long term, data samples for the suggested RIS-based hybrid precoding architecture can be produced via unsupervised learning.

To reduce the transmit power of the AP in an RIS-aided MISO setup with unpredictable channel circumstances, Lin *et al.* [102] developed a DRL technique with improved learning performance. The authors developed an optimization-driven DDPG method, which incorporates model-based optimization into the architecture of a model-free DDPG algorithm to aid information transfers from the AP to the single-antenna receiver. In a continuous action space, the DDPG method is used to address the optimization issues. The DDPG approach utilizes a DNN with a parameter to estimate the policy for the Q-value function. The DNN seeks to enhance the value function by actively adjusting the parametric policy in the gradient direction. The DDPG method can more efficiently lead to the discovery of combined beamforming, as demonstrated by the simulated data.

For maximizing the energy efficiency of non-orthogonal multiple access networks (NOMA), Liu *et al.* [94] explored the challenges of joint system implementation, power

distribution, determination of dynamic decoding order, and phase shift of RIS control in an RIS-aided system with multiple users while maintaining the data rate enhancement of the individual as required. The authors applied ML to solve this optimization challenge and suggested an LSTM-oriented echo state network algorithm based on an empirical dataset for anticipating the traffic demands in the case of an increased number of users in the future. Furthermore, a position-acquisition and phase-control method is proposed based on a decaying double DQN to identify the location of the RIS and control policy. By implementing a real dataset, an RNN is utilized to estimate the data traffic density. The BS serves as an agent in the DQN-based model. Because of having an installed controller, the BS can manage the phase shift, positioning of the RIS, and the policy of power allocation to the user. The energy efficiency of the RIS-NOMA combination was more significant than that of the benchmarks.

D. RESOURCE MANAGEMENT FOR RIS

Feng *et al.* [93] enhanced the model of RIS-aided phase shifts in downlink MISO communication networks to maximize the received signal-to-noise ratio (SNR). The authors used DRL because the level of examined resource allocation complexity is greater when constructing a realistic design of the phase-shifts solution. To avoid the constraints of the DQN, a DDPG-based technique was created to cope with continuous action spaces. During the initialization stage, four networks with uniformly distributed parameters were developed. In addition, with capacity, the experience of replay was constructed. At the start of every event, the phase shifts of all elements were randomly selected from 0 to 2π , without sacrificing its simplification. Both the critic and actor evaluation networks utilized the Adam optimizer for updating their parameters. Mathematical results show that a close to optimal signal-to-interference-plus-noise ratio (SINR) performance was achieved using the devised algorithm. Even though the performance of the model is more efficient than that of DQN, delayed convergence could be a limitation due to the broad parameters and divergence of the model from the desired position during initialization.

To combat strong propagating attenuations and enhance the transmission distance at THz-band frequencies, Huang *et al.* presented a multi-hop RIS-aided communication system, in which a DRL obtains the hybrid design of transmission beamforming at the BS and a phase-shift matrix [98]. As illustrated in Figure 16, the authors used numerous passive RISs to connect the BS to several single-antenna users. Two DNNs were used in the DRL structure. One DNN uses the critic network to assess the current policy associated with the rewards, while the other employs the actor network to estimate a policy based on the measured environment state and output of an action. According to the simulation findings, the proposed method can enhance the transmission distance of THz telecommunication by 50%.

Model aggregation for federated learning (FL) via radio channels is hampered by a lack of communication capacity.

Yang *et al.* [123] devised a concurrent access strategy aided by RIS to improve model aggregation performance, resulting in the creation of a connectivity-efficient FL framework for internet of things (IoT). The proposed approach comprises an FL system made up of 20 single-antenna IoT devices that are used to train a support vector machine (SVM) classifier using the Canadian Institute for Advanced Research (CIFAR-10) dataset that has been randomly distributed and divided. Lowering the model aggregation error, as measured by the MSE, is critical for improving the learning efficiency of the over-the-air computing-oriented FL. The channel environment between the aggregation and devices server is crucial for the MSE. With the help of the RIS, it is feasible to minimize the MSE of the model aggregation. RIS can obtain desirable channel responses by facilitating software-controlled phase shifts. The simulation results demonstrate that the RIS-enabled model aggregation produces a significantly greater convergence rate and considerably lower training loss. However, the computational complexity is quite high.

To provide good quality of network service for traffic moving through an obstructed area, Al-Hilo *et al.* [100] proposed a DRL framework with multi-binary action space to search a policy that maximizes the minimum average bit rate for vehicles using wireless scheduling. The authors suggested a route with no direct connectivity via a roadside unit (RSU). An obstacle is considered to block the line of sight. The RSU is expected to have a variety of channels to schedule the motor vehicles, and it must decide how to allocate its resources and adjust the RIS aspects if there are several vehicles on the road. Three linear layers are utilized for the DRL, with tanh as the activation function of the middle layer and softmax as the activation function of the output layer. Every internal layer has 64 units, and the Adam optimizer is used to reduce the loss function. The effectiveness of the solution technique was carefully examined by comparing it to other standards, and the framework was identified to be highly efficient in terms of the numerical findings. However, wireless resource allocation was not considered, where a spectrum can be provided to each vehicle depending on its unique requirements.

Ni *et al.* [104] jointly optimized the phase shifts and transmitted power while increasing the convergence rate of over-the-air FL (AirFL) to solve the device selection and undesired transmission error of FL in multiple RIS-aided systems. The suggested FL architecture was used to train a 7-layered CNN for image classification on the Modified National Institute of Standards and Technology dataset and a 50-layered residual network on CIFAR-10 dataset. The authors consider an RIS-assisted AirFL platform, IoT devices, and one BS, as shown in Figure 17, assuming that both the devices and the BS have a single antenna. The proposed approach accelerates the convergence rate and reduces aggregation error, according to the simulated data. However, this study did not include the privacy improvement strategy related to FL.

In [99], Kim *et al.* developed a dynamic control system based on multi-agent DRL that uses only index gradient

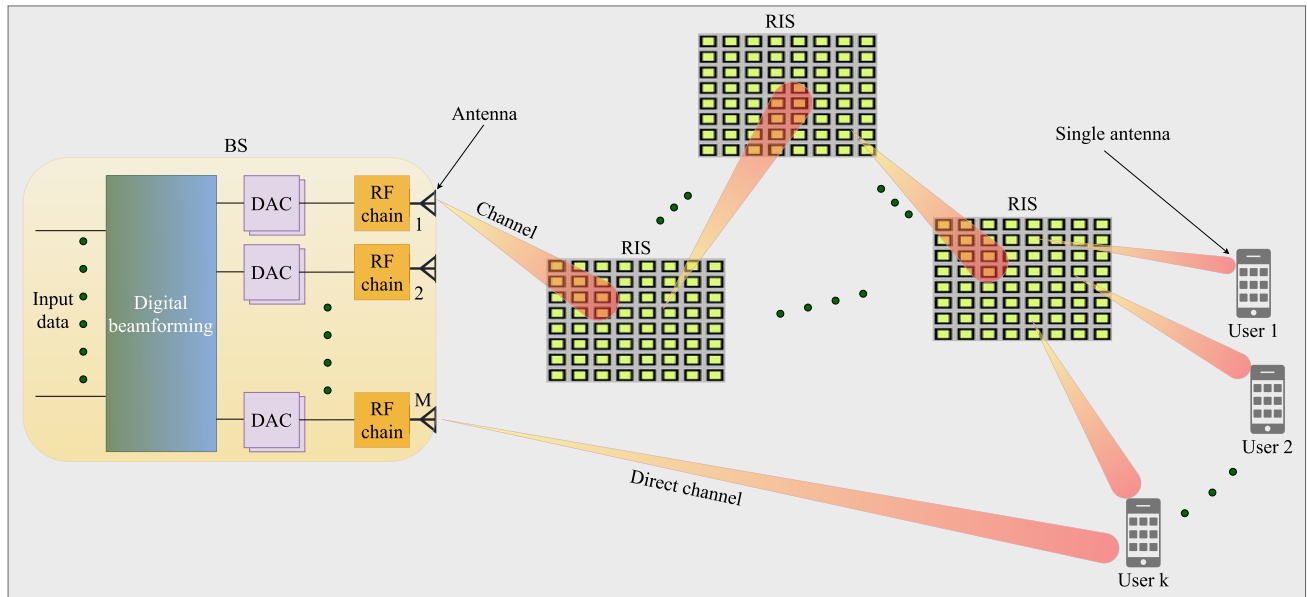


FIGURE 16. Multihop RIS system.

variables for the local RIS reflection beamformer to manage the localized user equipment broadcast powers and their combiners. The reward functions, action, and state represent the interdependencies in decision making at various BSs. For the uplink, the authors propose a multicell system with several RISs and cells, as shown in Figure 18. The system uses a DQN, where the BS has its DQN trained with weights. The suggested strategy achieves a significant improvement in the overall average rate in comparison with the baseline approaches, according to numerical data analysis.

To tackle a joint optimization issue for a NOMA downlink network using RIS, Yang *et al.* [97] provided a method based on the DDPG algorithm for constructing the phase shifts in RIS. The reward function was established by utilizing the sum rate of mobile clients; consequently, the best track of the agent was identified through the objective function. The phase-shift matrix of RIS defines the state space. In the learning phase, the action for every state is determined using an exponentially weighted algorithm, softmax, and a small-scale Rayleigh fading between the AP and clients. Figure 19, depicts the layout of the suggested system, in which the agent of the DDPG algorithm controls the reflecting components of the RIS and is sufficiently smart to learn the best phase shifts through exploitation and exploration. According to numerical data, the performance of the proposed framework may be enhanced by expanding the number of reflecting elements in the RIS and lowering the complexity of the RIS phase shifts.

Beam management (BM) is difficult for high-performance mmWave networks. To address this issue, Jia *et al.* [86] presented a DL-enabled BM system for RIS-aided mmWave networks, which analyzes the motion and environmental knowledge to achieve high system efficiency. The authors assume a standard RIS-aided mmWave infrastructure, in which

numerous RISs are installed that provide reliable and continuous connectivity for the intended location where the mmWave transmitted from the BS is hindered by the barriers. Whenever the real-time network information is fed into the DNN model, the ideal network parameters may be forecasted instantly without the need for complex optimization processes, thereby substantially reducing the system overhead. Figure 20 depicts the time frame model of RIS-aided mmWave platform, which includes data transmission, channel acquisition, beam monitoring, and beam training during initial access. The DL-enabled BM framework achieves improved signal quality and changeover success rate, according to the simulation findings. However, this study does not provide BM for multicell networks.

To solve the phase shift and combined trajectory design issue, in [103] the authors proposed a decaying DQN (D-DQN)-based algorithm to predict the phase shift of the RIS and trajectory of the unmanned aerial vehicle (UAV) while ensuring that data demands of the client are satisfied in the process. The control center, which governs both the UAV and the RIS, operates as an agent in the D-DQN based method. A UAV is used to provide wireless services to several single-antenna users. Throughout an area, all users are considered to be roaming. It is assumed that an RIS is installed on the wall of a tall building with reflecting components to improve the quality of cellular connectivity by establishing a cascaded virtual line-of-sight propagation between the clients and the UAV. The MDP is characterized by the states, environment, reward function, and actions. The cycle is formed by the state transition function. The transitions of a new state occur after one MDP cycle, based on the actions performed and the previous state. In comparison to the traditional DQN method, the suggested D-DQN oriented approach uses a decaying learning rate to achieve a balance between the training speed

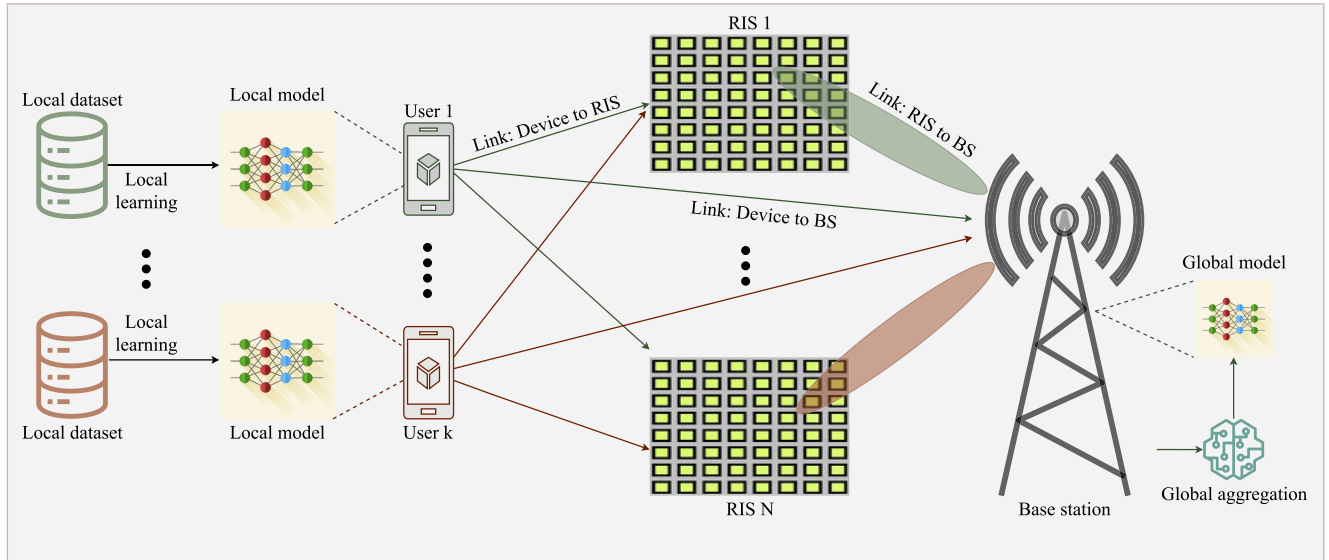


FIGURE 17. FL in multiple RIS system.

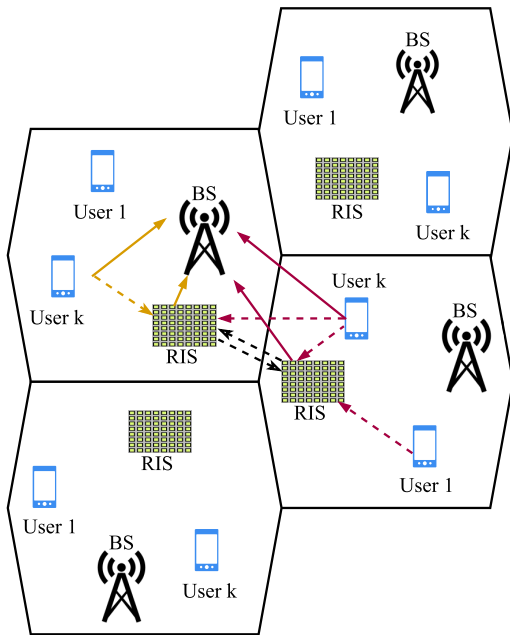


FIGURE 18. Multicell system.

and oscillation. However, the motion of the UAV in terms of speed variation was not included in this study.

Obtaining the activation pattern of the IoT devices ahead of the UAV flight is a complicated task. To address this issue, Samir *et al.* [101] developed a DRL-focused PPO to shift the phase of RIS components to detect the unpredictability of IoT device activation patterns, and performed connectivity planning to reduce the predicted cumulative age-of-information and regulate the height of the UAV. When the transmission power of IoT devices is increased, the SNR obtained at the BS improves directly. Nevertheless, in certain IoT applications, increasing the quantity of reflecting components per

RIS increases the obtained age of information and SNR by improving the quality of the communication link between the BS and the IoT devices. However, no investigation has been performed in the case where the source or destination nodes have several antennas.

E. DETECTION FOR RIS

In [79], the authors utilized a DL-based method to estimate and identify symbols in an RIS-aided wireless system. To calculate the phase angles and channels of a reflected received signal by an RIS, a fully connected NN was used. This allows symbol identification by not the application of any specialized adjustment in the obtained pilot signal, which substantially lowers the overhead that is necessary for the RIS-aided network. The bit error rate (BER) of the system was enhanced using DL. The proposed DeepRIS network consists of three hidden layers. The model was trained over several iterations so that it could be robust to data overfitting. Moreover, DeepRIS outperformed the traditional MMSE and least squares estimators. However, the necessity of gathering a large amount of data for the diverse user positions complicates the proposed technique.

To cope with the signal interference that deteriorates the SINR of the RIS, Yang *et al.* [75] demonstrated that CNN can be used as part of a traditional RIS controller to recognize the conflicting devices using incident signals. To enhance the wireless connection quality, the authors examined an RIS-aided uplink network composed of users, a group of RISs, and a BS. Although each RIS functions as a nearly passive surface, the RIS controller may consume energy in the given model. To build the spectrum-learning approach, online CNN inference, offline CNN training, and acquisition of RF traces are the three key components. Convolutional layers with ReLU activation functions are used in the trained CNN

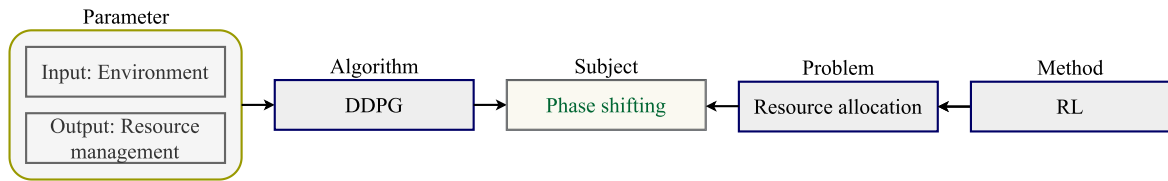


FIGURE 19. Use of deep deterministic policy gradient in the phase shift of RIS.

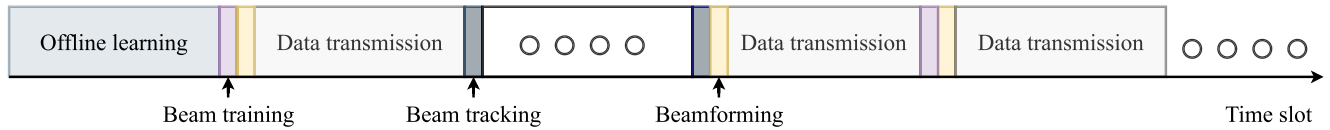


FIGURE 20. Time frame model for beam management.

model, which are accompanied by two fully connected layers. The ability of DL to increase performance was validated by simulation findings. However, the CNN needs to be retrained if the distribution of RF data changes considerably.

The authors suggested a feasible ML approach for wireless fingerprinting localization in RIS-aided settings using a NN with ReLu as the activation function and a single hidden layer of size 100 [81]. The proposed system model consists of a receiver, an RIS, and a transmitter and assumes that the AP and RIS are both linked to a service provider who could also manage the RIS settings. Figure 21 illustrates the suggested method in which every RIS setting is viewed as a feature, and the goal is to choose the best collection of features from several options. If features are chosen negligently, the feature set may comprise redundant, irrelevant, and inaccurate information. The simulation results demonstrate that a supervised learning-oriented approach in RIS for selecting features can improve the detection performance and minimize the location collection time. However, the work does not include different scenarios or the use of multiple RISs.

Vaca-Rubio *et al.* presented a computer vision method based on SVM and TL to analyze radio graphics created by the RIS to identify abnormalities along the path of a robot [124]. The authors analyzed an industrial situation in which a robot maintains a predetermined route that assumes that it could divert from the intended path and pursue an undesirable direction owing to random factors. While the target device travels along the path, the training data are acquired by sampling the received power at different temporal moments. The proposed model contains a VGG19 structure, and the last fully linked layer is eliminated during the modification process. The findings demonstrate that RIS-aided detection provides considerable accuracy and appears to have a wide range of applications in indoor industrial settings.

F. SECURITY

While the key generation rate (KGR) is typically restricted by wireless channel dynamics, it can also be hindered by other

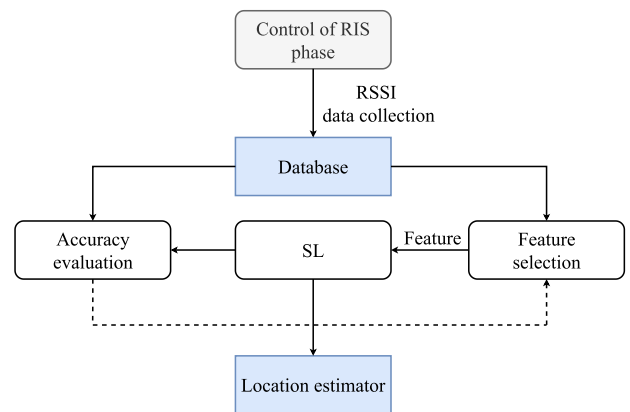


FIGURE 21. Representation of localization algorithm.

factors. In wireless fading of channels, there is randomness that affects the secrecy of the produced key. Weak probability extraction from the channel limits the KGR in static or slow-fading scenarios. To enhance the KGR of the physical layer key generation on the basis of CSI, Jiao *et al.* [106] presented an interactive quantization level forecasting model using the ML technique. The authors considered a wireless system with two single-antenna users that were aided by RIS. By exploring and retrieving unpredictability from bi-directional channels, users communicate in the TDD mode to generate a secret key. A function is used to track fresh channel observational data in real time, and the suggested method assesses the training data at the beginning. The FNN used in the model consists of two hidden layers, where a tansig function for the hidden layer and purelin function for the output layer, is used as an activation function. When the SNR is high, the forecasting model prefers to use high quantization levels to decrease the bit disagreement ratio, whereas low quantization levels are used when the SNR is lower to maintain the bit disagreement rate at a minimum.

Yang *et al.* [95] proposed a secure beamforming strategy with DRL in the RIS-assisted communication network for physical layer security, where adjustment of the reflecting

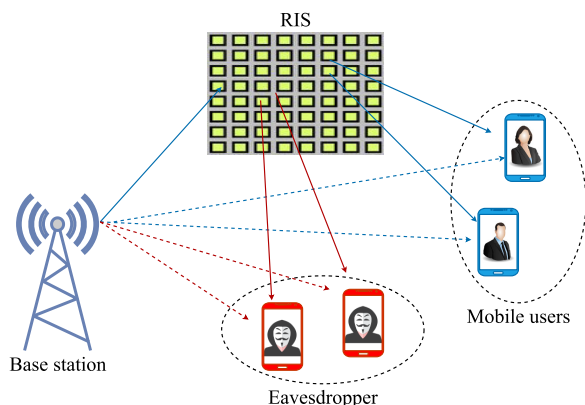


FIGURE 22. Secure communication for RIS.

elements in RIS is performed to ensure the safe communication of numerous genuine clients in the presence of numerous spies, as shown in Figure 22. A model for the joint optimization of beamforming in the RIS and BS is proposed to increase the secrecy rate of the system under the premise of quality, service needs, and time-varying channel conditions. A prioritized experience replay and post-decision state are used to improve secrecy and learning efficiency. During the training steps, the controller using RL modifies the parameters of the network and monitors the state of the present system, which includes several factors such as total CSI for all clients, data rate of transmission, and the expected secrecy rate. Then, the DQN is fed with the vectors of the state to train the model. The ϵ -greedy policy is used for balancing exploitation and exploration, while a random action is determined depending on the knowledge gained from the environment. The RL agent obtains a reward from the environment after completing the selected action and observes the state shift. Simulation findings show that the proposed secure beamforming strategy improves the system secrecy with a good probability. However, the model incorporates complex computations and large amounts of data; consequently, the additional computational power will affect its efficiency.

Li *et al.* proposed a privacy-preserving ML-boosted communication systems by adding FL in the RIS-aided wireless technology to solve the privacy breach of user information [125]. Multiple RIS-assisted IoT network transmission and single RIS-aided outdoor connectivity strategies were proposed by the authors. FL-based RIS-aided outdoor communications will utilize distributed learning to train the best DNN model for projecting user channels to the best RIS configuration matrix and achieving high-speed wireless communication while protecting privacy. The FL is used in IoT network connectivity to enhance several RISs in parallel under the protection of the private CSI method, allowing the best possible rate of the combined signal, which would be the convergence of signals across all RISs. According to the simulation results, the proposed architecture improves user privacy while maximizing the attainable rate of the receiver. However, owing to the common wireless links utilized by the

transmitted information, the efficiency of global aggregation may suffer.

The concern of privacy issues in RIS-aided connectivity was addressed by Ma *et al.* [105] by suggesting an FL-based algorithm for achieving high-speed communication while maintaining the privacy of an individual. A local model can be trained in FL using the local dataset of each participant, and a global model was constructed by combining all local models. The starting configuration of each device in the next training cycle is retrieved from the global model. To reach the ideal global model, all steps are iterated until the global model output converges. For FL to understand the mapping function between the RIS configuration matrix and the CSI, a DNN is adopted. As per the input CSI, the DNN output is determined by the optimal attainable rate of the device. A transmitter is assumed to interact with an RIS-aided receiver via a server linked to the RIS for data processing. The performance of the proposed algorithm using FL may effectively satisfy the theoretical value and be greater than 90% of that obtained by the centralized ML while safeguarding user privacy, according to the simulation findings. However, the utilization of a CNN can be a better choice because it may result in greater performance; moreover, because of the simple RIS design, connecting a complex parameter server may be problematic. For model training, it is more convenient to use the BS to handle the parameter server.

VI. CHALLENGES AND FUTURE RESEARCH

Owing to its ability to enhance resource usage, RIS has garnered considerable interest. The RIS in wireless networks shows considerable potential, but a variety of difficulties and limitations must be addressed, as shown in Figure 23.

A. ISSUES WITH THE CHANNEL

Ongoing studies on channel fading in RIS have focused on simpler wireless link designs. Fading schemes similar to traditional phased arrays and multi-antenna systems are often used. Accurate modeling is required for the transmission of signals dispersed by metasurfaces in assessing the performance boundaries in wireless connections using RISs. Assuming that the inter-antenna length is more significant than half the wavelength, these approaches could be suitable for a large number of cheap antennas, while their applicability to meta-atom based RIS requires additional research. Furthermore, sub-wavelength structured fading models must be developed at both the visible and microscopic levels for easy incorporation into communication schemes.

RISs do not require amplifiers. Therefore, the critical question is how it can notify the channel condition to the receiver, transmitter, or any controller responsible for determining the optimal phases and assessing the channels necessary for phase optimization. Consequently, it might be beneficial to incorporate an energy harvesting unit in the RIS to drive low-powered sensors that monitor the channels and present the results to a gateway for sending these results back to the

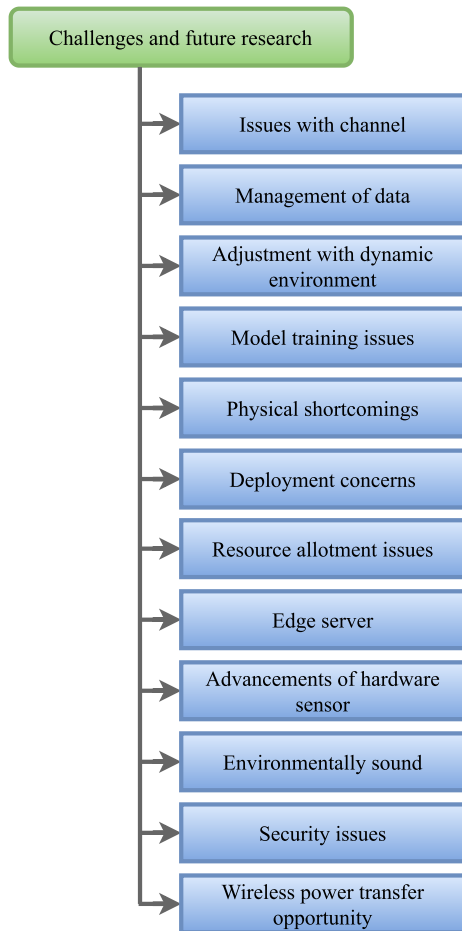


FIGURE 23. Categorization of challenges and future research.

network controller. Thus, RIS-based communication can be a part of an energy-efficient technology.

A high proportion of recent research advancements have concluded that there is an unbroken CSI accessible at the BS. Even so, because of the passive nature of RIS, the exchange of CSI and acquisition is not an easy task to perform. Quick and precise data gathering is critical in RIS communication systems; however, significant training is required to achieve this.

Unresolved areas of research include the evaluation of technical limitations and study of the attainable performance of RIS in frequency-selective channel fading. Deploying an RIS in a near-field environment may have several intriguing uses and possible advantages. Currently, supervised DL is widely utilized to solve channel problems to maximize efficiency, and DL techniques may be used to explore CSI structures that are more than just linear correlations.

B. MANAGEMENT OF DATA

Data frameworks depend on the information, which raises important issues from the perspective of data control management, in which stability and latency impose limitations on the feedback loop. Analyzing and modeling RIS-powered

wireless communications is often more complicated than existing communication schemes. To overcome the complexities of such systems, data-driven approaches based on learning algorithms provide new opportunities. Channel sensing may gather significant data owing to their sensing potential and availability of large scattering components, which opens the door to data-driven DL methods. An adequate dataset is required for the DL model to obtain good training results. In the compressive sensing method, a large dataset is not required. However, in other disciplines, obtaining adequate training data may be problematic. It is possible that if only limited data are obtained for estimating the parameters in the NN, we will end up with considerable variation and overfitting. A need for more investigation on information storage solutions is required to enable the administration of data with a responsive and rapid system.

C. ADJUSTMENT WITH DYNAMIC ENVIRONMENT

User mobility varies significantly in several areas and at different times. The variable factors of a user are in terms of speed, direction, acceleration, and angle. It is difficult to learn the traffic pattern of a user because of the mobile and dynamic natures of the traffic obtained from the client. Furthermore, the employed parameters cannot be defined as a constant value owing to the limitless quantity of mobility. Obtaining labels is a significant challenge in supervised learning in the existing method; therefore, other learning solutions such as computer vision (using CNNs) and neural language processing (based on LSTM) techniques may be suitable for assessing and understanding the traffic model.

D. MODEL TRAINING ISSUES

Training is a key component in the application of ML in an RIS. Consequently, it is vital to understand the performance limitations of training models under dynamic conditions. The equilibrium between the development of non-static settings and training rate is unclear. Some of the existing DL methods still have several practical issues. However, standard ML methods must not be disregarded, and the situations in which classical approaches are preferable to DL techniques must be identified. Although DL is a potential technique for RIS, it requires considerable effort and a large amount of data to obtain the necessary results. To further study the convergence of performance, an improvement in the training rate could be achieved at the possible cost of loss in performance.

Hyperparameter adjustment is required for ML model training [126]. The effectiveness of ML model is determined by how well it is trained using data. It is not easy to train a model with a large amount of data. Minor adjustments in parameter values can have a significant impact on the performance of the ML model. Furthermore, the training of the model can be viewed as a computationally demanding operation that uses considerable CPU and GPU resources, especially for deep networks. However, it is challenging for the BS server to handle this massive quantity of data on its own, in addition to distinguishing between the relevance of

the data among all the information necessary to run RIS. It is critical to utilize the essential data and reject the irrelevant data for the server to decrease the workload of the server and allow it to use the relevant data when learning is taking place.

E. PHYSICAL SHORTCOMINGS

Meta-surfaces are composed of subwavelength objects with complex formations. Consequently, the absence of precise and manageable models that characterize customizable meta-surfaces as a factor of their EM characteristics is a key constraint in the ongoing studies on RIS. In most of the studies, it is assumed that the metasurface behaves like an element of reflection. However, it is crucial to bear in mind that meta-surfaces are not just created to reflect the wave; they are also designed to perform other tasks because their reaction to radio waves is affected by elements such as materials of composition, polarization, and angles. The spatial pairing between the antenna components is generally ignored while modeling the RIS. To determine whether the performance of a metasurface-based RIS increases with the denser packing of antenna elements, suitable uniform designs must be used. For metasurface architectures, EM-based circuit designs may be employed, which clearly considers mutual linkage and the configurations among the unit cells. Thus, we can effectively overcome the sub-wavelength barrier and provide possible advantages of metasurface constructions.

F. DEPLOYMENT CONCERNS

Wireless networks driven by the RIS have a wide range of possible uses, both outdoors and indoors. Deployment is one of the most important design considerations when incorporating an RIS into a communication system. In RIS-assisted connectivity, the proper deployment of the RIS can enhance the system to ensure that the receiver, transmitter, and RIS have a line of sight. Hence, the number and location of RIS deployment are essential factors that must be resolved.

Large amounts of bandwidth are accessible in the mmWave frequency. Therefore, mmWave connectivity is capable of delivering gigabits per second information speed, which is essential for higher data rate applications. However, because of the short waveforms, it suffers from extreme blockage and imposes substantially greater energy usage with higher equipment costs because of the increased number of active antennas that are running at significantly higher frequencies when compared to lower wireless frequency systems. Owing to the intelligent reflections that RIS provides, short-wave-related difficulties can be effectively handled by accurately deploying RISs in mmWave platforms to construct a virtual line-of-sight channel between the users and the BS that can combat the obstructions between them.

G. RESOURCE ALLOTMENT ISSUES

A metasurface functioning as an anomalous reflector must be modeled using the principles of physics, as the phase and

amplitude interactions are not self-governing. The reactive nature of the metasurface architecture allows the incorporation of physics-based models to study the influence of the RIS sub-wavelength structural system, design limitations, and radio wave manipulation characteristics. In vehicle models, the spectrum may be allocated to every vehicle depending on its specific demands. Therefore, the RIS phase-shift setup technique may be modified to understand the impact of different link qualities for each individual vehicle based on the given wireless capacity. Researchers have considered that RIS installations at permanent sites can only support nearby users. To further enhance the productivity of RIS with UAVs, integrated wireless communication can be deployed, and the speed of the UAV must be optimized.

H. EDGE SERVER

In many ways, edge computing is similar to cloud computing. Owing to the passive beamforming capabilities of RIS, mobile edge computing devices may use RISs to increase the potential of connecting with the edge server and transferring computationally intensive jobs to the edge server to reduce delay rather than experiencing excessive transmission power usage. When using model iteration instead of collecting user data, FL may be used to boost edge computing and enhance privacy. To conduct global control smartly, FL will work in conjunction with edge devices, which are locally trained, and then transfer model parameters to the control center. More attention should be focused on latency and privacy problems in this area.

I. ADVANCEMENTS OF HARDWARE SENSOR

It would be fascinating to investigate methods for building entirely freestanding RIS designs, in which the RIS is not controlled by the infrastructure but will operate autonomously while interacting with the surroundings. Many mobile devices may now be fitted with a variety of sensors owing to the advent of sensor technology. In the configuration of beam control, visual data may be derived from previous knowledge. Incorporating visual data with other sensor data may be used to provide motion information and enhance awareness of the surroundings for the mmWave network that is combined with an RIS.

J. ENVIRONMENTALLY SOUND

An advantage of the RIS is that it can manipulate EM waves without requiring power amplifiers and other high-power demanding equipment; therefore, we position RIS at the forefront of the eco-friendly device list. The benefits of these concepts are that they can make metasurfaces more recyclable and help reduce the exposure of people to EM fields. However, it complicates the process of obtaining the essential environmental parameters for configuring and optimizing the RIS. Hence, it is crucial to properly evaluate the essential exchange required between RIS power usage and operating complications.

K. SECURITY ISSUES

Security and privacy are equally key issues in future RIS-based initiatives because of their vulnerable surroundings. Consequently, there has been an upsurge in the creation of encrypted data transfer methods that rely on the physical features of communication. To avoid possible data leaks and maintain data security, RIS offers a path to manipulate the propagation environment around unsafe endpoints. With numerous eavesdroppers and authorized users in a place where RIS will be densely installed, it is important to enhance network secrecy. Using the FL technique, noise may be used to improve connection security by decreasing privacy leaks and making the fields of research worth exploring.

L. WIRELESS POWER TRANSFER OPPORTUNITY

To extend the life of the battery of IoT devices in next-generation cellular networks that link billions of limited power-consuming nodes, wireless power exchange can be a viable technique. Interactions between devices are more dispersed and varied than traditional downstream broadcasts from a multi-antenna transmitter to receivers, thus posing additional challenges for RIS-aided communications. Diverse approaches have been presented to mitigate significant energy losses through vast distances to improve the reliability of power transfer. For wireless power transmission, the advantages of RIS beamforming are highly dependent upon the availability of channel information at the transmitter, which is gained at the expense of energy and time; therefore, improper training will lead to unreliable channel information, which consumes considerable energy at the receiving side and also leaves less opportunity for the collection of energy.

VII. CONCLUSION

In this paper, we provide a comprehensive review of the emergence of ML in RIS advancements. We started with an overview of RIS and explained the implementation of ML in RIS along with its limitations. Finally, the potential obstacles and open research problems when using ML in RIS have been identified. Major technological issues should be overcome to adequately address the considerable architectural demands in future networks. Over the next few years, the implementation of RIS in mapping, propagation, localization, signal processing, and resource allocation idea is expected to achieve revolutionary results. Further research on RIS and ML can be conducted on various topics that could have a significant impact in the field of wireless communication for the next generation of networks.

REFERENCES

- [1] (2021). E. Perspectives and C. Report. *Cisco Annual Internet Report—Cisco Annual Internet Report (2018–2023) White Paper*. Accessed: Dec. 28, 2021. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectiv%es/annual-internet-report/white-paper-c11-741490.html>
- [2] M. A. El Mossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 990–1002, Sep. 2020.
- [3] B. Sheen, J. Yang, X. Feng, and M. M. U. Chowdhury, "A deep learning based modeling of reconfigurable intelligent surface assisted wireless communications for phase shift configuration," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 262–272, 2021.
- [4] C.-X. Wang, M. D. Renzo, S. Stanczak, S. Wang, and E. G. Larsson, "Artificial intelligence enabled wireless networking for 5G and beyond: Recent advances and future challenges," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 16–23, Feb. 2020.
- [5] X. Liu, M. Chen, Y. Liu, Y. Chen, S. Cui, and L. Hanzo, "Artificial intelligence aided next-generation networks relying on UAVs," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 120–127, Feb. 2021.
- [6] Q.-V. Pham, N. T. Nguyen, L. B. Le, and K. Lee, "Intelligent radio signal processing: A contemporary survey," 2020, *arXiv:2008.08264*.
- [7] N. T. Nguyen, K. Lee, and H. DaiIEEE, "Application of deep learning to sphere decoding for large MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6787–6803, Oct. 2021.
- [8] L. V. Nguyen, D. H. N. Nguyen, and A. L. Swindlehurst, "DNN-based detectors for massive MIMO systems with low-resolution ADCs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.
- [9] S. Sun, Z. Cao, H. Zhu, and J. Zhao, "A survey of optimization methods from a machine learning perspective," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3668–3681, Nov. 2020.
- [10] K. M. Faisal and W. Choi, "A study on machine learning-based approaches for reconfigurable intelligent surface," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2021, pp. 227–232.
- [11] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.
- [12] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, and J. De Rosny, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–20, May 2019.
- [13] Y.-C. Liang, R. Long, Q. Zhang, J. Chen, H. V. Cheng, and H. Guo, "Large intelligent surface/antennas (LISA): Making reflective radios smart," *J. Commun. Inf. Netw.*, vol. 4, no. 2, pp. 40–50, Jun. 2019.
- [14] E. Basar and I. Yildirim, "Reconfigurable intelligent surfaces for future wireless networks: A channel modeling perspective," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 108–114, Jun. 2021.
- [15] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.
- [16] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "Realizing wireless communication through software-defined HyperSurface environments," in *Proc. IEEE 19th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2018, pp. 14–15.
- [17] G. Lavigne, K. Achouri, V. S. Asadchy, S. A. Tretyakov, and C. Caloz, "Susceptibility derivation and experimental demonstration of refracting metasurfaces without spurious diffraction," *IEEE Trans. Antennas Propag.*, vol. 66, no. 3, pp. 1321–1330, Mar. 2018.
- [18] S. V. Hum and J. Perruisseau-Carrier, "Reconfigurable reflectarrays and array lenses for dynamic antenna beam control: A review," *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 183–198, Jan. 2014.
- [19] J. Huang, Q. Li, Q. Zhang, G. Zhang, and J. Qin, "Relay beamforming for amplify- and-forward multi-antenna relay networks with energy harvesting constraint," *IEEE Signal Process. Lett.*, vol. 21, no. 4, pp. 454–458, Apr. 2014.
- [20] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosy, M. Fink, S. Tretyakov, and S. Shamai, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, 2020.
- [21] F. Yang and Y. Rahmat-Samii, *Surface Electromagnetics: With Applications in Antenna, Microwave, and Optical Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2019.
- [22] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
- [23] S. Tretyakov, *Analytical Modeling in Applied Electromagnetics*. Norwood, MA, USA: Artech House, 2003.

- [24] E. F. Kuester, M. A. Mohamed, M. Piket-May, and C. L. Holloway, "Averaged transition conditions for electromagnetic fields at a metafilm," *IEEE Trans. Antennas Propag.*, vol. 51, no. 10, pp. 2641–2651, Oct. 2003.
- [25] A. Epstein and G. V. Eleftheriades, "Huygens' metasurfaces via the equivalence principle: Design and applications," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 33, no. 2, pp. A31–A50, 2016.
- [26] T. Saeed, C. Skitsas, D. Kouzapas, M. Lestas, V. Soteriou, A. Philippou, S. Abadal, C. Liaskos, L. Petrou, J. Georgiou, and A. Pitsillides, "Fault adaptive routing in metasurface controller networks," in *Proc. 11th Int. Workshop Netw. Chip Architectures (NoCArc)*, Oct. 2018, pp. 1–6.
- [27] A. C. Tasolamprou, M. S. Mirmoosa, O. Tsilipakos, A. Ptilakakis, F. Liu, S. Abadal, A. Cabellos-Aparicio, E. Alarcon, C. Liaskos, N. V. Kantartzis, S. Tretyakov, M. Kafesaki, E. N. Economou, and C. M. Soukoulis, "Intercell wireless communication in software-defined metasurfaces," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.
- [28] H.-T. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces: Physics and applications," *Rep. Prog. Phys.*, vol. 79, no. 7, Jun. 2016, Art. no. 076401.
- [29] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah, "Holographic mimo surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.
- [30] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Reconfigurable intelligent surfaces: Three myths and two critical questions," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 90–96, Dec. 2020.
- [31] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.
- [32] V. Cherkassky and F. M. Mulier, *Learning From Data: Concepts, Theory, and Methods*. Hoboken, NJ, USA: Wiley, 2007.
- [33] T. M. Mitchell, *The Discipline of Machine Learning*. Pittsburgh, PA, USA: Carnegie Mellon Univ., School of Computer Science, 2006.
- [34] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.
- [35] C. Rudin and K. L. Wagstaff, "Machine learning for science and society," *Mach. Learn.*, vol. 95, pp. 1–9, Apr. 2014.
- [36] Y. Anzai, *Pattern Recognition and Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2012.
- [37] B. Adam and I. F. Smith, "Reinforcement learning for structural control," *J. Comput. Civil Eng.*, vol. 22, no. 2, pp. 133–139, 2008.
- [38] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 160, pp. 3–24, Oct. 2007.
- [39] J. H. Friedman, R. Tibshirani, and T. Hastie, *The Elements of Statistical Learning*, vol. 1, no. 10. New York, NY, USA: Springer, 2001.
- [40] M. Ponsen, M. E. Taylor, and K. Tuyls, "Abstraction and generalization in reinforcement learning: A summary and framework," in *Proc. Int. Workshop Adapt. Learn. Agents*, Budapest, Hungary, 2009, pp. 1–32.
- [41] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.
- [42] P. Beyens, M. Peeters, K. Steenhaut, and A. Nowe, "Routing with compression in wireless sensor networks: A Q-learning approach," in *Proc. 5th Eur. Workshop Adapt. Agents Multi-Agent Syst. (AAMAS)*, Paris, France, 2005.
- [43] J. A. Boyan and M. L. Littman, "Packet routing in dynamically changing networks: A reinforcement learning approach," in *Proc. Adv. Neural Inf. Process. Syst.*, 1994, pp. 671–678.
- [44] R. Sun, S. Tatsumi, and G. Zhao, "Q-MAP: A novel multicast routing method in wireless ad hoc networks with multiagent reinforcement learning," in *Proc. IEEE Region 10th Conf. Comput., Commun., Control Power Eng.*, vol. 1, Oct. 2002, pp. 667–670.
- [45] A. Mukherjee, V. Keshary, K. Pandya, N. Dey, and S. C. Satapathy, "Flying ad hoc networks: A comprehensive survey," in *Information and Decision Sciences*. Singapore: Springer, 2018, pp. 569–580.
- [46] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [exploratory DSP]," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 145–154, Jan. 2011.
- [47] D. C. Cireşan, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Deep, big, simple neural nets for handwritten digit recognition," *Neural Comput.*, vol. 22, no. 12, pp. 3207–3220, 2010.
- [48] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.
- [49] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 1, pp. 30–42, Jan. 2011.
- [50] X.-W. Chen and X. Lin "Big data deep learning: Challenges and perspectives," *IEEE Access*, vol. 2, pp. 514–525, 2014.
- [51] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*. Hershey, PA, USA: IGI Global, 2010, pp. 242–264.
- [52] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.
- [53] V. Francois-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," 2018, *arXiv:1811.12560*.
- [54] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017.
- [55] D. Bau, J.-Y. Zhu, H. Strobelt, A. Lapedriza, B. Zhou, and A. Torralba, "Understanding the role of individual units in a deep neural network," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 48, pp. 30071–30078, Dec. 2020.
- [56] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time-series," in *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA, USA: MIT Press, 1998.
- [57] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, p. 1097.
- [58] C. Li, Y. Wu, X. Yuan, Z. Sun, W. Wang, X. Li, and L. Gong, "Detection and defense of DDoS attack-based on deep learning in OpenFlow-based SDN," *Int. J. Commun. Syst.*, vol. 31, no. 5, p. e3497, Mar. 2018.
- [59] A. Ahmad, S. Ismail, and D. Samaon, "Recurrent neural network with backpropagation through time for speech recognition," in *Proc. IEEE Int. Symp. Commun. Inf. Technol. (ISCIT)*, vol. 1, Oct. 2004, pp. 98–102.
- [60] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [61] X. Li and X. Wu, "Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 4520–4524.
- [62] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 30, no. 1, 2016, pp. 1–7.
- [63] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," 2015, *arXiv:1509.02971*.
- [64] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.
- [65] W. Qingqing and Z. Rui, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2019.
- [66] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2283–2314, Jun. 2020.
- [67] H. Gacanin and M. Di Renzo, "Wireless 2.0: Toward an intelligent radio environment empowered by reconfigurable meta-surfaces and artificial intelligence," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 74–82, Dec. 2020.
- [68] A. M. Elbir and K. V. Mishra, "A survey of deep learning architectures for intelligent reflecting surfaces," 2020, *arXiv:2009.02540*.
- [69] J. Wang, W. Tang, Y. Han, S. Jin, X. Li, C.-K. Wen, Q. Cheng, and T. J. Cui, "Interplay between RIS and AI in wireless communications: Fundamentals, architectures, applications, and open research problems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2271–2288, Aug. 2021.
- [70] S. Zhang, M. Li, M. Jian, Y. Zhao, and F. Gao, "AIRIS: Artificial intelligence enhanced signal processing in reconfigurable intelligent surface communications," 2021, *arXiv:2106.00171*.
- [71] W. Tang, M. Z. Chen, J. Y. Dai, Y. Zeng, X. Zhao, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 180–187, Apr. 2020.
- [72] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

- [73] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021.
- [74] X. Yuan, Y.-J. A. Zhang, Y. Shi, W. Yan, and H. Liu, "Reconfigurable-intelligent-surface empowered wireless communications: Challenges and opportunities," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 136–143, Apr. 2021.
- [75] B. Yang, X. Cao, C. Huang, C. Yuen, L. Qian, and M. D. Renzo, "Intelligent spectrum learning for wireless networks with reconfigurable intelligent surfaces," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3920–3925, Apr. 2021.
- [76] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *IEEE Access*, vol. 9, pp. 44304–44321, 2021.
- [77] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M.-S. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9223–9228, Aug. 2020.
- [78] A. M. Elbir, A. Papazafiroopoulos, P. Kourtessis, and S. Chatzinotas, "Deep channel learning for large intelligent surfaces aided mm-Wave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1447–1451, Sep. 2020.
- [79] S. Khan, K. S. Khan, N. Haider, and S. Y. Shin, "Deep-learning-aided detection for reconfigurable intelligent surfaces," 2019, *arXiv:1910.09136*.
- [80] C. Huang, G. C. Alexandropoulos, C. Yuen, and M. Debbah, "Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [81] C. Ly Nguyen, O. Georgiou, and G. Gradoni, "Reconfigurable intelligent surfaces and machine learning for wireless fingerprinting localization," 2020, *arXiv:2010.03251*.
- [82] M. Xu, S. Zhang, C. Zhong, J. Ma, and O. A. Dobre, "Ordinary differential equation-based CNN for channel extrapolation over RIS-assisted communication," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 1921–1925, Jun. 2021.
- [83] S. Zhang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Deep learning optimized sparse antenna activation for reconfigurable intelligent surface assisted communication," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6691–6705, Oct. 2021.
- [84] Y. Lu, M. Hao, and R. MacKenzie, "Reconfigurable intelligent surface based hybrid precoding for THz communications," 2020, *arXiv:2012.06261*.
- [85] N. Abuzainab, M. Alrabeiah, A. Alkhateeb, and Y. E. Sagduyu, "Deep learning for THz drones with flying intelligent surfaces: Beam and hand-off prediction," 2021, *arXiv:2102.11222*.
- [86] C. Jia, H. Gao, N. Chen, and Y. He, "Machine learning empowered beam management for intelligent reflecting surface assisted mmWave networks," *China Commun.*, vol. 17, no. 10, pp. 100–114, Oct. 2020.
- [87] M. A. Aygul, M. Nazzal, and H. Arslan, "Deep learning-based optimal RIS interaction exploiting previously sampled channel correlations," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.
- [88] N. T. Nguyen, L. V. Nguyen, T. Huynh-The, D. H. N. Nguyen, A. L. Swindlehurst, and M. Juntti, "Machine learning-based reconfigurable intelligent surface-aided MIMO systems," 2021, *arXiv:2105.00347*.
- [89] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, "Unsupervised learning for passive beamforming," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1052–1056, May 2020.
- [90] H. Song, M. Zhang, J. Gao, and C. Zhong, "Unsupervised learning-based joint active and passive beamforming design for reconfigurable intelligent surfaces aided wireless networks," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 892–896, Mar. 2021.
- [91] C. Huang, R. Mo, and Y. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Jun. 2020.
- [92] A. Taha, Y. Zhang, F. B. Mismar, and A. Alkhateeb, "Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020, pp. 1–5.
- [93] K. Feng, Q. Wang, X. Li, and C. Wen, "Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 745–749, May 2020.
- [94] X. Liu, Y. Liu, Y. Chen, and H. V. Poor, "RIS enhanced massive non-orthogonal multiple access networks: Deployment and passive beamforming design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1057–1071, Apr. 2021.
- [95] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 375–388, Jan. 2021.
- [96] Q. Zhang, W. Saad, and M. Bennis, "Millimeter wave communications with an intelligent reflector: Performance optimization and distributional reinforcement learning," *IEEE Trans. Wireless Commun.*, early access, Sep. 3, 2021, doi: [10.1109/TWC.2021.3107520](https://doi.org/10.1109/TWC.2021.3107520).
- [97] Z. Yang, Y. Liu, Y. Chen, and N. Al-Dhahir, "Machine learning for user partitioning and phase shifters design in RIS-aided NOMA networks," 2021, *arXiv:2101.01212*.
- [98] C. Huang, Z. Yang, G. C. Alexandropoulos, K. Xiong, L. Wei, C. Yuen, and Z. Zhang, "Hybrid beamforming for RIS-empowered multi-hop terahertz communications: A DRL-based method," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2020, pp. 1–6.
- [99] J. Kim, S. Hosseinipour, T. Kim, D. J. Love, and C. G. Brinton, "Multi-IRS-assisted multi-cell uplink MIMO communications under imperfect CSI: A deep reinforcement learning approach," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–7.
- [100] A. Al-Hilo, M. Samir, M. Elhatab, C. Assi, and S. Sharafeddine, "Reconfigurable intelligent surface enabled vehicular communication: Joint user scheduling and passive beamforming," 2021, *arXiv:2101.12247*.
- [101] M. Samir, M. Elhatab, C. Assi, S. Sharafeddine, and A. Ghayeb, "Optimizing age of information through aerial reconfigurable intelligent surfaces: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3978–3983, Apr. 2021.
- [102] J. Lin, Y. Zout, X. Dong, S. Gong, D. T. Hoang, and D. Niyato, "Deep reinforcement learning for robust beamforming in IRS-assisted wireless communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2020, pp. 1–6.
- [103] X. Liu, Y. Liu, and Y. Chen, "Machine learning empowered trajectory and passive beamforming design in UAV-RIS wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 2042–2055, Jul. 2021.
- [104] W. Ni, Y. Liu, Z. Yang, H. Tian, and X. Shen, "Federated learning in multi-RIS aided systems," 2020, *arXiv:2010.13333*.
- [105] D. Ma, L. Li, H. Ren, D. Wang, X. Li, and Z. Han, "Distributed rate optimization for intelligent reflecting surface with federated learning," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.
- [106] L. Jiao, G. Sun, J. Le, and K. Zeng, "Machine learning-assisted wireless phy key generation with reconfigurable intelligent surfaces," in *Proc. 3rd ACM Workshop Wireless Secur. Mach. Learn.*, 2021, pp. 61–66.
- [107] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," 2013, *arXiv:1312.5602*.
- [108] A. M. Elbir, K. V. Mishra, and Y. C. Eldar, "Cognitive radar antenna selection via deep learning," *IET Radar, Sonar Navigat.*, vol. 13, no. 6, pp. 871–880, Jun. 2019.
- [109] A. M. Elbir, "CNN-based precoder and combiner design in mmWave MIMO systems," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1240–1243, May 2019.
- [110] A. M. Elbir and K. V. Mishra, "Joint antenna selection and hybrid beamformer design using unquantized and quantized deep learning networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1677–1688, Mar. 2020.
- [111] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [112] L. Subrt and P. Pechac, "Intelligent walls as autonomous parts of smart indoor environments," *IET Commun.*, vol. 6, no. 8, pp. 1004–1010, May 2012.
- [113] A. M. Elbir and S. Coleri, "Federated learning for channel estimation in conventional and RIS-assisted massive MIMO," 2020, *arXiv:2008.10846*.

- [114] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 898–912, Feb. 2022.
- [115] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.
- [116] A. Alkhateeb, "DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications," 2019, *arXiv:1902.06435*.
- [117] T. Lin and Y. Zhu, "Beamforming design for large-scale antenna arrays using deep learning," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 103–107, Jan. 2020.
- [118] W. Lee, M. Kim, and D.-H. Cho, "Transmit power control using deep neural network for underlay Device-to-Device communication," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 141–144, Feb. 2019.
- [119] C. Liaskos, A. Tsioliaridou, S. Nie, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "An interpretable neural network for configuring programmable wireless environments," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [120] S. Dörmer, S. Cammerer, J. Hoydis, and S. ten Brink, "Deep learning based communication over the air," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 132–143, Feb. 2018.
- [121] S. Gong, J. Lin, J. Zhang, D. Niyato, D. I. Kim, and M. Guizani, "Optimization-driven machine learning for intelligent reflecting surfaces assisted wireless networks," 2020, *arXiv:2008.12938*.
- [122] G. Lee, M. Jung, A. T. Z. Kaskari, W. Saad, and M. Bennis, "Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces," in *Proc. ICC - IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [123] K. Yang, Y. Shi, Y. Zhou, Z. Yang, L. Fu, and W. Chen, "Federated machine learning for intelligent IoT via reconfigurable intelligent surface," *IEEE Netw.*, vol. 34, no. 5, pp. 16–22, Sep. 2020.
- [124] C. J. Vaca-Rubio, P. Ramirez-Espinosa, R. Jess Williams, K. Kansanen, Z.-H. Tan, E. de Carvalho, and P. Popovski, "A primer on large intelligent surface (LIS) for wireless sensing in an industrial setting," 2020, *arXiv:2006.06563*.
- [125] L. Li, D. Ma, H. Ren, D. Wang, X. Tang, W. Liang, and T. Bai, "Enhanced reconfigurable intelligent surface assisted mmWave communication: A federated learning approach," *China Commun.*, vol. 17, no. 10, pp. 115–128, Oct. 2020.
- [126] G. Luo, "A review of automatic selection methods for machine learning algorithms and hyper-parameter values," *Netw. Model. Anal. Health Informat. Bioinf.*, vol. 5, no. 1, pp. 1–16, Dec. 2016.



K. M. FAISAL (Graduate Student Member, IEEE) received the B.S. degree from the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh, in 2019. He is currently pursuing the M.Sc. degree with the Smart Networking Laboratory, Chosun University, Gwangju, Republic of Korea. His current research interests include reconfigurable intelligent surfaces, heterogeneous wireless networks, and machine learning-based optimization.



WOOYEOL CHOI (Member, IEEE) received the B.S. degree from the Department of Computer Science and Engineering, Pusan National University, Busan, Republic of Korea, in 2008, and the M.S. and Ph.D. degrees from the School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea, in 2010 and 2015, respectively. He was a Senior Research Scientist with the Korea Institute of Ocean Science and Technology (KIOST), Ansan, Republic of Korea, from 2015 to 2017, and a Senior Researcher with the Korea Aerospace Research Institute (KARI), Daejeon, Republic of Korea, from 2017 to 2018. He is currently an Assistant Professor with the Department of Computer Engineering, Chosun University, Gwangju. His research interests include cross-layer protocol design, learning-based resource optimization, and experiment-driven evaluation of wireless networks.

• • •