Received 15 April 2024; revised 2 June 2024; accepted 6 June 2024. Date of publication 11 June 2024; date of current version 6 August 2024.

Digital Object Identifier 10.1109/OJAP.2024.3412609

A State-of-the-Art Survey on Advanced Electromagnetic Design: A Machine-Learning Perspective

MASOUD SALMANI ARANI (Graduate Student Member, IEEE), REZA SHAHIDI (Senior Member, IEEE), AND LIHONG ZHANG (Senior Member, IEEE)

Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

CORRESPONDING AUTHOR: M. S. ARANI (e-mail: msalmaniaran@mun.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant RGPIN-2020-07155 and Grant RGPIN-2024-04448; in part by the Canada Foundation for Innovation (CFI) under Grant 40646; and in part by Memorial University of Newfoundland.

ABSTRACT Research on electromagnetic (EM) components is essential to enabling the design and optimization of such devices as antennas and filters, leading to improved functionality, reduced costs, and enhanced overall performance. This paper presents an overview of recent developments in optimization and design automation techniques for EM-component design and modeling. Limitations of conventional optimization methods are discussed, while the need for novel machine learning techniques capable of handling multiple objectives and large design spaces is highlighted. In this study, existing methods in the literature are reviewed from four viewpoints: structural view, algorithm view, component view, and application view. Different schemes in distinct design stages or applications are examined with advantages and drawbacks laid out for easier comprehension. Finally, to broaden the scope of optimization in the field of EM design and modeling, some prospective trends are pointed out to shed light on emerging research hotspots.

INDEX TERMS Optimization, design automation, deep neural networks, inverse modeling, microwave computer-aided design (CAD).

I. INTRODUCTION

ELECTROMAGNETIC (EM) components are well-known for their nonlinear behaviors and vulnerability to the environment. As a result, modeling, simulation, and verification of complex EM designs are becoming increasingly complicated. These problems may become even more challenging when a high-dimensional input space (design space) is required to form a desired multi-objective output space (performance space). To delve into solutions for demanding specifications, optimization methods and design automation are good fits, especially in the wake of recent advances in machine learning (ML) [1].

While optimizations for EM-based problems have a vital role in designing microwave or millimeter wave (mmWave) components, they suffer from great computational complexities in practice. This is mainly due to the repetitive EM simulations needed to investigate the impacts of different geometrical parameters as well as technology and/or material coefficients, which typically serve as design variables [2]. Thus, data acquisition and data sampling are of great importance in guiding the optimization process toward the optimum solutions and reducing the computational burden.

In EM component design, optimization problems frequently arise aiming at multiple objectives in the context of numerous variables [3]. While mathematical programming approaches can prove to offer a single optimal solution if the problem is convex, population-based evolutionary algorithms (EAs) excel in finding a set of optimal solutions by balancing conflicting objectives for any (including non-convex) problems [4]. Nevertheless, conventional EAs tend to converge at a slower pace compared to the mathematical programming methods due to exponential growth of the design space with the increasing number of optimization variables. This phenomenon is commonly referred to as the curse of dimensionality [5].

In recent years, much effort has been devoted to developing optimization and design automation algorithms to mitigate the challenges associated with error-prone operations and to avoid time-consuming manual design tasks. The significant benefits of design automation include higher speed, lower human endeavor and errors, and higher productivity. Electronic design automation (EDA) tools, such as Ansys HFSS, CST and FeKo, have been employed in the design of various electromagnetic structures. Although such computer-aided design (CAD) tools are getting more mature for the field of electromagnetics, they are typically used for design entry, simulation, and rendering purposes [6]. Three literature reviews in [7], [8], [9] have explored the application of ML in antenna design and optimization, focusing on initial ML methods. However, when dealing with the real-world EM design challenges, there are typically numerous variables to adjust and various constraints to meet throughout the optimization process. As a result, the demand has arisen for more robust and expressive models to effectively address these complexity challenges.

Yu et al. [10] provided a review of the state-of-the-art AI-assisted methodologies only for microwave filter design. Moreover, an excellent overview of artificial neural network (ANN) techniques for microwave CAD design was presented in [11]. However, the limited scope of ANN tends to constrain the authors from further investigating more advanced ML-based methods, such as reinforcement learning. In a similar work [12], a survey of ML techniques was carried out for the modeling of radio wave propagation. Most recently, Cha et al. [13] utilized natural language processing (NLP) and ML techniques to analyze the research of antennas and propagation (A&P) through extensive unstructured data drawn from openly published scientific papers and patents, forecasting future A&P research trends.

In modern AI research, various architectures have been developed to enhance the efficiency of ANNs. For instance, transformers utilize a self-attention mechanism to manage sequential data, excelling in capturing long-range dependencies and revolutionizing various fields like NLP. As another example, autoencoders complement these architectures by enabling efficient data compression and reconstruction, providing robust feature extraction capabilities. Additionally, neural networks with transfer function (neuro-TF), a blend of neural networks and traditional transfer functions, are used effectively in parameterized EM modeling. By leveraging knowledge of transfer functions, such as poles and zeros, this method simplifies highly nonlinear problems into more manageable ones for neural networks to learn, particularly in scenarios with sharp resonances versus frequency. Together, these techniques enhance the adaptability and efficiency of neural networks in handling diverse tasks. We will explore the applicability of these methods in EM design and optimization in more detail in subsequent sections.

Design and optimization methods for electromagnetic components may be categorized in different ways. To better understand the main contributions and advantages of existing techniques and then envision future developments, in this paper we opt to categorize them from four viewpoints: 1) structural view, 2) algorithm view, 3) component view, and 4) application view. As shown in Fig. 1, we begin with the structural view, examining essential structural elements within the typical process of EM optimizations. Next, the algorithm view explores computational strategies, including machine learning and heuristic algorithms. After that, the component view examines specific EM components such as antennas and metamaterials, discussing unique challenges and optimization strategies. Finally, the application view showcases real-world benefits of machine learning techniques across diverse fields of applications. These perspectives offer a comprehensive understanding of current and future directions in design and optimization of EM components, including antennas, filters and metamaterials.

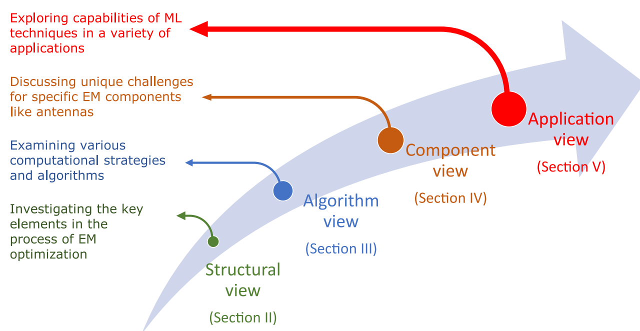


FIGURE 1. Top-level organization of the paper.

The rest of this review paper is organized as follows. We first look at the key components in the general structure of the EM optimizations in Section II. Then different algorithms employed in a variety of EM design areas are presented in Section III. The component view is presented in Section IV, while various applications of EM optimization algorithms are discussed in Section V. Some discussions on literature statistics and extended insights into future research & development are offered in Section VI and finally, Section VII presents the main conclusions of this paper.

II. STRUCTURAL VIEW

During the optimization process, it is possible to substitute computationally-expensive simulations with inexpensive surrogate model predictions [14]. Consequently, the optimization time can be significantly reduced. Regarding the availability of a model to aid in the optimization process, there are two different approaches, namely offline surrogate modeling and online surrogate modeling. The former scheme uses a fixed dataset to construct a surrogate model, which remains static and does not adapt to new data. In contrast, the online surrogate modeling scheme continuously updates the surrogate model in real time based on new observations, allowing for dynamic adaptation and improved accuracy while the optimization progresses. Here we investigate three main factors for online surrogate modeling as shown in Fig. 2, which are model itself, search operator, and framework.

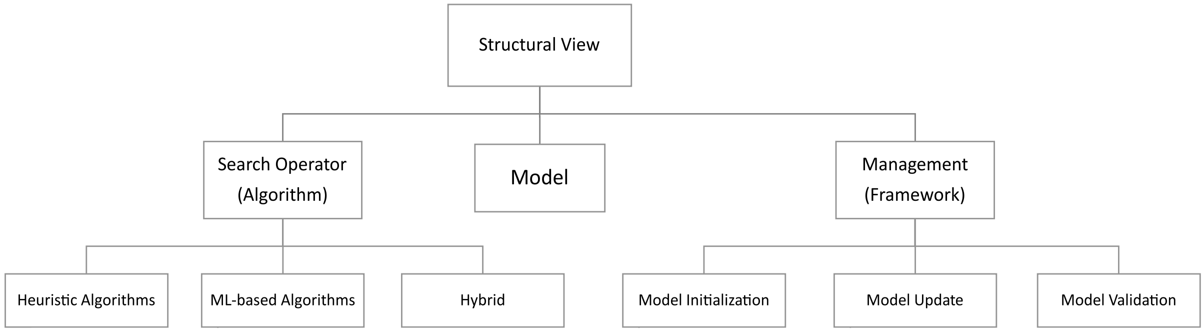


FIGURE 2. Classification from the structural view.

A. MODEL

A model is a representation of a real-world phenomenon or complex system from which various types of engineering tasks including design optimization and sensitivity analysis can be carried out. Hence, the design process can be formulated as an optimization problem [10]

$$x^* = \arg\min_x U(R(x)),\quad s.t.\quad x \in [x_{LB}, x_{UB}] \tag{1}$$

where x represents a set of design parameters, and x^* indicates the optimal design parameters. $U(\cdot)$ stands for the objective function to be optimized, while $R(x)$ denotes the EM responses. In addition, the range of search space is denoted as $[x_{LB}, x_{UB}]$.

Nonetheless, use of extensive and accurate models for exploring the best designs is often impractical due to the overwhelming computation required. To expedite the search process, an effective approach is to employ surrogate models, which are also referred to as *approximation models* formulated by

$$x^* = \arg\min_x R_s(U(R(x))), \tag{2}$$

where $R_S(\cdot)$ represents the response values obtained from a constructed surrogate model. These models act as representations of the original models, enabling them to replace costly simulation models by approximating their input-output responses [14].

Among the surrogate models, Gaussian process (GP) model has been widely employed by the virtue of its great ability to learn and reduced data requirements for initialization. However, updating the GP model during each iteration within the optimization process is still a highly time-consuming task. Liu et al. [15] proposed a method for complex antenna design, which integrated an adaptive GP and radial basis function (RBF) model with differential evolution (DE) operations to build a framework to enable many design candidates to share the same GP model when possible. However, the training time of the GP model undergoes cubic expansion as the number of training data points, design variables and specifications increases. As an enhancement to the previous work, a Bayesian neural network (BNN) model with lower computational complexity and higher prediction uncertainty was employed in [16] to replace the GP model as the surrogate model. This improvement surpasses the previous work, reducing the total optimization time by over 50%. Nevertheless, by shifting to a different design requirement, there is no avenue to reuse the previous results or data, which might be a direction for further improvement.

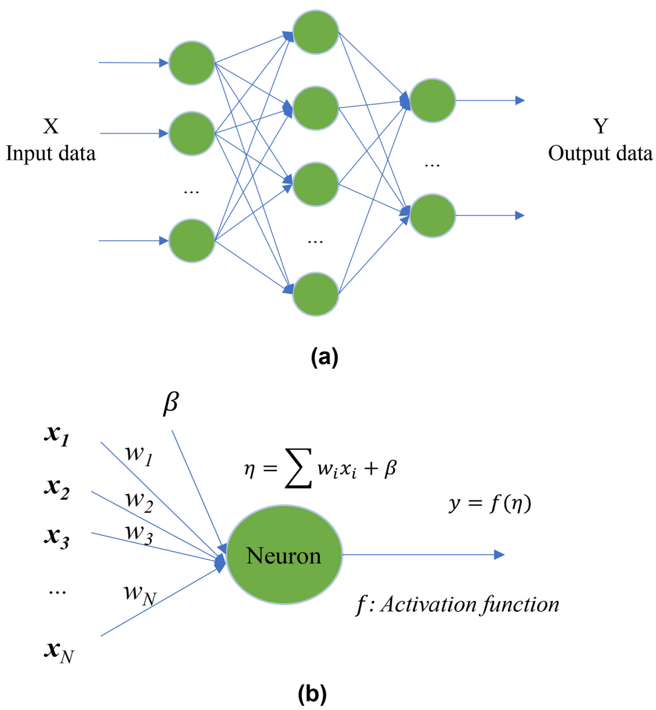


FIGURE 3. Structure of the ANN: (a) three-layer fully-connected network and (b) a neuron with the details adapted from [10].

ANNs, as another popular model, have the potential to serve as a machine learning model which can perform regression to effectively map intricate, high-dimensional, and nonlinear functional connections between input and output parameter spaces. This makes ANNs a useful resource that can support computational electromagnetic (CEM) approaches in the optimization and modeling of intricate static electromagnetic design challenges [11]. The graphical description of an ANN is illustrated in Fig. 3(a), where

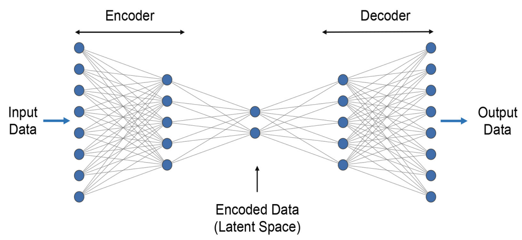


FIGURE 4. The architecture of an autoencoder.

the known data X is fed into the input layer of the neural network, which then makes predictions Y as its outputs. In addition, a single neuron (also called *perceptron* in ML) is shown in more detail in Fig. 3(b), where a single neuron function involves three key components: weights, bias and activation function. Weights $\{w_1, w_2, \ldots, w_N\}$ are the parameters that the neuron uses to scale its input signals $\{x_1, x_2, \ldots, x_N\}$. Each input signal is multiplied by its corresponding weight, and these products are summed up to produce the weighted sum. Then a bias β is added to this weighted sum to obtain a linear expression η in the inputs. Finally, the activation function f takes η as input and typically applies a nonlinear transformation to derive the output. The activation function often introduces nonlinearity into the model, enabling it to learn and represent complex patterns within the data. Some of the well-known examples for the activation functions include Sigmoid, ReLU, Tanh and Softmax, each with its own specifications and suitability for different tasks.

Convolutional neural networks (CNNs), which are widely known for their remarkable capabilities for image processing and feature extraction, are a class of ANN models. As an example, in [17] Shibata et al. proposed a convolution-autoencoder method that utilized unsupervised learning to extract geometric features from a dataset of planar bandpass filter (BPF) images without relying on any EM simulation outputs. As shown in Fig. 4, the autoencoder compresses the input data to a lower dimensional space to capture its principal components, then decodes it back to the original size. This process reduces noise and perturbations in the data. After training a CNN, the S-parameters could be rapidly calculated simply by giving an image of BPF circuit pattern as an input to the trained CNN. However, in their work only four design variables were considered in highly-restricted ranges. While the dataset consisted of fewer than 7,000 entries, it might be adequate for some problems with narrow design scope, but may fall short for complex problems.

Introduced in [18], the space mapping (SM) technique has been further applied to various engineering applications. The concept of space mapping relies on the presence of both coarse and fine models. The coarse models, such as equivalent circuit models, are typically computationally efficient but less accurate. Conversely, the fine models offer accurate solutions but require significant computational resources. The SM technique establishes a mathematical relationship between the fine and coarse models, enabling the combination between the fine-model accuracy and coarse-model efficiency [19]. In [20], Melgarejo et al. employed SM techniques to tune microwave filters at X-band frequencies with different coarse-model fidelities. Its major drawback is that the accuracy of the solution is highly dependent on the quality of the surrogate model used as the coarse model. In other words, constructing a surrogate model that accurately captures the behavior of the original problem may be a challenging task, requiring significant computational resources and expertise [21].

Instead of considering the entire response, feature-based optimization (FBO) methods, also known as *response feature methods*, use a set of suitably selected characteristic points of the system outputs to find a highly linear relationship in the frequency domain. For instance, in the case of antenna design, these feature points might be the antenna resonant frequencies or the points corresponding to S_{11} at the level of -10dB. In [22], a feature-based approach was employed to enhance the input tolerance of an antenna design, while remaining in the feasible design space. In general, the process of defining and extracting features from EM-simulated antenna responses, as well as defining the objective function based on these features, needs to be customized for specific types of antenna responses and design tasks. This limitation presents a significant drawback of the feature-based optimization technique in terms of automation, as it requires a certain level of user experience and interaction to set up the optimization framework. In another work [23], Pietrenko-Dabrowska and Koziel proposed a unified definition of response features so that the characteristic points can be defined and extracted automatically regardless of a specific set of performance specifications, and subsequently used to formulate a feature-based objective function.

B. SEARCH OPERATOR

A search operator, also known as *optimization algorithm*, is a computational method or procedure used to find the optimal solution to a problem within a given search space. It is a key component in the optimization process, where the goal is to find the best possible solution that satisfies certain criteria or constraints. There are numerous types of search operators such as genetic algorithms (GAs), differential evolution (DE), Bayesian optimization, etc., each with their own characteristics and suitability for different types of problems.

To further accelerate the search within the design space, some dedicated techniques are commonly used. One of these is the trust region (TR) methods, which were employed in [24] to limit the search space of design variables to further improve the convergence rate of optimization. Moreover, neuro-TF has recently been developed to create parametric models of EM responses. The neuro-TF method, as shown in Fig. 5, leverages a transfer function to capture the highly nonlinear relationship between EM responses and frequency. This simplifies the task for an ANN by reducing the

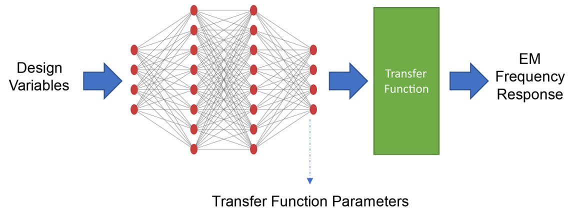


FIGURE 5. Structure of the neuro-TF method.

nonlinearity in the relationships between transfer function parameters and design variables. Consequently, the neuro-TF method is more accurate and robust. The trained neuro-TF model serves as a surrogate, enabling rapid surrogate-based EM design optimization. For instance, in [25] Zhuo et al. proposed a method that introduces a polynomial function and a pole-residue based transfer function to represent EM responses. Initially, the entire EM response was represented by the pole-residue based transfer function, which was then divided into multiple sub-transfer functions, each consisting of a pole and its corresponding residue. A new algorithm for discerning smoothness was also introduced to evaluate the smoothness of each sub-transfer function response and separate the pole-residue pairs whose sub-transfer function responses were determined to be smooth. To explore more features of optimization algorithms in depth, we will dedicate Section III to the algorithm view.

C. FRAMEWORK/MODEL MANAGEMENT

Model management methods refer to the techniques and practices used to effectively handle and control surrogate models in the context of modeling and optimization tasks. They involve various processes, strategies, and considerations to ensure the surrogate model to accurately represent the underlying system and contribute to the optimization process. Three pivotal tasks are essential to model management as shown in Fig. 2: model initialization, update, and validation.

One of the challenges in model initialization is adequate data availability, a significant concern when developing ANN models. In simple terms, the accuracy of an ANN model improves as more data samples become available. However, obtaining a large quantity of data samples is unavoidably time-consuming and requires substantial resources, such as 3-D EM simulations. To alleviate this issue, some studies have been carried out to accelerate the conventional process of solving Maxwell's equations. We will further discuss state-of-the-art data generation techniques in Section V-A.

For model management, another important consideration is the process of updating the model, which involves selecting the next data point to be utilized for improvement. In particular, for EM modeling where structures are often complex, selection of the next step in the optimization process greatly impacts on the overall time required for optimization. For example, Bayesian optimization strategies are defined in an indirect manner by optimizing what is known as an acquisition function. This function evaluates and assigns a score to possible observation paths. This score reflects how much these locations are believed to enhance the optimization process and guides the optimizer to select the next sampling point [26]. Many variations of the acquisition function concept have been developed to expedite the model update process. In [24], Zhou et al. applied a modified acquisition function, which is the product of the classical expected improvement (EI) function and a penalty function to produce multiple update points for the parallel computation of EM responses at each iteration cycle. However, due to the nature of Bayesian optimization, this method is computationally expensive and sensitive to initial conditions, with limited scalability to high-dimensional problems.

Model validation, as another consideration in model management, refers to the process of assessing and evaluating performance, accuracy, and generalizability of a machine learning or statistical model. It involves testing the model using independent data that was not used in the model training phase. The purpose of model validation is to determine how well the model can predict or estimate outcomes on unseen data. Mean square error (MSE) is a commonly utilized measurement in assessing the performance of models [27].

III. ALGORITHM VIEW

To meet ever-increasingly demanding design requirements and aggressive development turnaround, conventional trial-and-error approaches are being replaced by global and local optimization methods. Minimizing the difference between the simulated and expected objective function amounts is the goal of optimization algorithms. Numerical methods, such as the finite-element method (FEM), method of moments (MoM) and finite difference methods (FDM), are mostly used in CAD software to carry out a series of EM simulations. Nevertheless, these methods come at a huge cost of memory and runtime [28].

Besides that, in most cases the objective functions to be optimized are challenging to evaluate. They may have many local optima and the derivative information is usually unavailable. That is to say, the widely-used gradient-based methods tend to become inadequate, and thus the need for a global optimizer arises with the goal of more effectively searching the design parameter space [28]. Although there are numerous statistical methods and optimization algorithms available in the literature, this survey aims to identify the most prominent ones in the context of EM component design. Fig. 6 summarizes different EM optimization methods from the algorithmic perspective. At the highest level, there are three core categories, namely heuristic, ML-based, and hybrid methods. The heuristic methods can be further subdivided into two major categories, EAs, and homotopy methods. In addition, the ML-based algorithms can be broken down into supervised, unsupervised, and reinforcement learning methods. Finally, the hybrid methods are the ones, which apply multiple optimization methods together to find the problem solution.

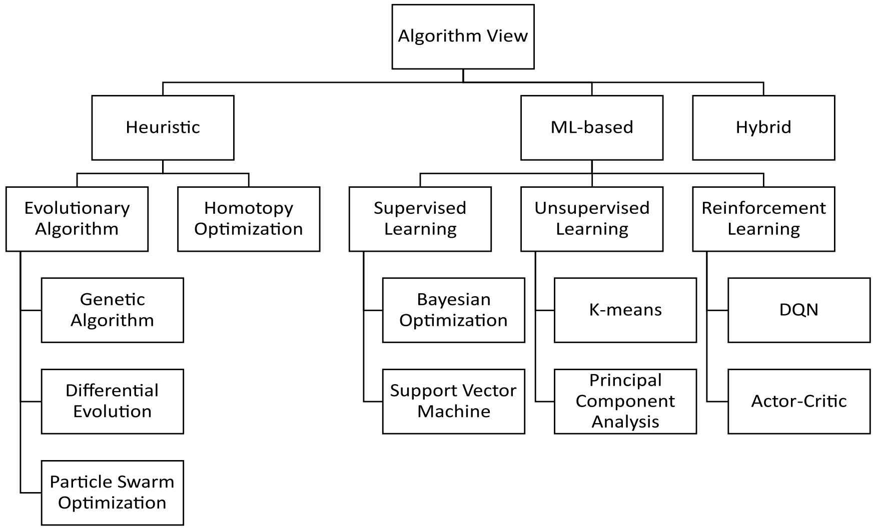


FIGURE 6. Categorization of different optimization methods from the algorithm view.

A. HEURISTIC METHODS

A heuristic optimization algorithm is a problem-solving technique used to find approximate solutions to complex optimization problems. Heuristic algorithms draw inspiration from natural or human problem-solving strategies, intuition, and experience, rather than relying on rigorous mathematical proofs or exhaustive search procedures.

One of the most important categories of global optimization techniques is EA, which is based on the idea of natural evolution. In these algorithms, a group of candidate solutions, called *population*, is evolved within the design space to find the global optimum. In the past decades, population-based iterative algorithms such as GAs [29], [30], [31], [32], [33], particle swarm optimization (PSO) [34], [35], [36], DE [37], [38] and grey wolf optimization (GWO) [39], [40], [41] have attracted significant interest. However, these methods typically suffer from high computational cost, which basically stems from the necessity of running the simulation iteratively and then making performance predictions for various combinations of input values [14], [42].

It is always important to choose a set of initial design parameter values close enough to the optimum in order to find a satisfactory result. Otherwise, we may fail to approach the optimal with regard to the desired performance criteria. Even though in some cases the initial parameters are significantly distant from the optimal solution, Zhao and Wu utilized homotopy optimization (HO) to search for an optimum solution for microwave and millimeter-wave filter design [43]. The homotopy method utilizes the principles of topology to produce a sequence of solutions for nonlinear systems by incorporating a parameter for controlling convergence. Instead of directly solving the desired problem, this approach formulates a series of nonlinear optimization problems. Moreover, Roy and Wu [44] employed HO to collect the data from different cruciform coupler designs to train an ANN model. After the development of this ANN model, the circuit parameters for various specifications of a cruciform coupler could be obtained without a need for repetitive utilization of an optimization algorithm. Notwithstanding the benefit of being able to handle non-convex optimization problems, this method suffers from high computational cost and excessive sensitivity to the choice of homotopy parameters.

B. ML-BASED METHODS

In general, ML methods refer to the algorithms with the capability of learning patterns and features to reduce human involvement and increase the predictability of an unknown event. They might be trained by some already-existing dataset or generated on purpose, without any prior knowledge of their environment. Over the past decade, there has been significant progress in developing high-expressive models, especially neural networks with many hidden layers, known as *deep neural network models*. This advancement has reignited interest in machine learning algorithms within the research community. Deep neural networks have been leveraged for extended capacity to model complex relationships within large data samples, and eventually, to tackle challenges in high-dimensional modeling and optimization problems. In Fig. 7, an EM application example is illustrated for each category of ML methods. In the sequel, we will elaborate on various ML algorithms and their respective categorization and applications.

1) SUPERVISED LEARNING METHODS

In supervised learning methods, for each element of the input, the corresponding output, called *label*, is provided to help the model to be trained to recognize meaningful patterns [45]. The precision of this class of ML method is highly-dependent on the amount of data presented as input

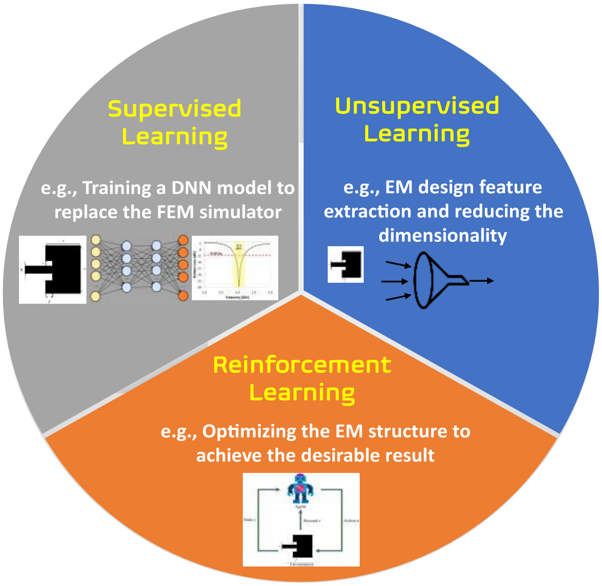


FIGURE 7. Categorization of machine learning (ML) techniques.

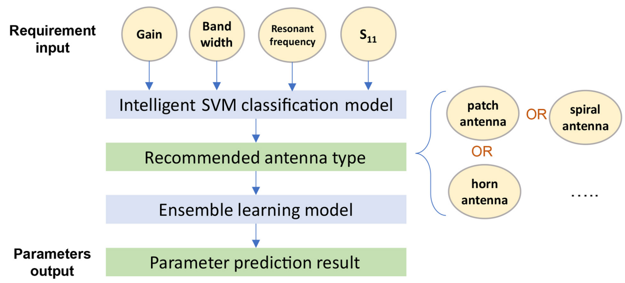


FIGURE 8. Intelligent antenna synthesis framework proposed in [47].

to the training process. This is the major drawback of the supervised learning methods.

One of the most widely-used supervised learning methods is Bayesian optimization (BO), in which the prior information provided by the optimization history is used to explore the design space more efficiently. By employing this approach, the BO framework is made up of two major parts. The first is a Gaussian process (GP) model playing the role of surrogate model. The other is the acquisition function or loss function [46]. For instance, Zhou et al. [24] used a GP model along with parallel EM simulations by using a modified acquisition function to obtain multiple updated points in a dynamic sampling design space. Nevertheless, not much significant enhancement in simulation runtime and iterations can be achieved by using this approach compared to more conventional methods, such as SADEA and pure BO.

As another popular supervised learning method, support vector machines (SVMs) are used for both classification and regression tasks. They work by finding the optimal hyperplane in a high-dimensional space that best separates data points belonging to different classes. Linear SVM is employed when data points can be readily separated by a straight line or hyperplane. On the other hand, nonlinear SVM is used when the data cannot be effectively separated using a straight line/hyperplane. In such cases, nonlinear SVM utilizes kernel functions to transform the data from a nonlinear space to a linear space. This transformation enables the data to be classified by finding an appropriate hyperplane in the transformed space. According to the antenna performance of interest, Shi et al. proposed an intelligent antenna synthesis method, in which the antenna type is automatically determined based on the performance specifications using an SVM classifier as shown in Fig. 8 [47]. Then its optimal geometric values are achieved through an ensemble learning model, which is a cooperative decision-making mechanism that constructs and integrates multiple learners to accomplish the learning process.

2) UNSUPERVISED LEARNING METHOD

In contrast to supervised learning methods, in unsupervised learning methods there is no need to use labeled data to train the model. Feature recognition and finding similarities for data clustering are the major benefits of unsupervised learning methods to help classification or data extraction.

One popular clustering algorithm is the K-means algorithm, which aims to partition a set of observations into K clusters. Here each observation belongs to a cluster with reference to the nearest centroid. Nevertheless, the primary issues with K-means are its sensitivity to the initial selection of cluster centers, as well as the requirement of knowing the number of classes a priori. The K-means algorithm seeks to minimize the sum of squared distances between observations and the centroids of their assigned clusters. However, since the algorithm may converge to a local minimum, multiple initializations with different starting centroids may be necessary to find the global minimum [48]. In this regard, in [49] Zhang et al. proposed a K-means algorithm to speed up the optimization process for designing a MIMO antenna with decoupling elements. The population was divided into multiple clusters using the K-means method, followed by the selection of an appropriate mutation strategy for each cluster based on its average fitness. Their results show a runtime saving of at least 28% compared to the GA, PSO, and DE algorithms.

As one of the most widely used multivariate techniques in statistics, principal component analysis (PCA) is an unsupervised learning method, whose main purpose is to examine the underlying structure of a set of variables and the covariance or correlation structure between them for dimensionality reduction. This technique is commonly utilized to reduce the complexity of data and reveal its underlying patterns without a need for predefined labels or categories [50], [51]. By reducing the dimensionality of data, PCA can help to compress large datasets and extract the most important features. In [52], Sedaghat et al. proposed a technique called PCA-based data compression to identify and eliminate redundant information from the response data in the training set of inverse modeling of microwave filters. This study presented a novel approach to determine the optimal number of PCA coefficients, which corresponds to the optimal compression level. By selecting the appropriate dimension for the compressed input space using orthonormal PCA coefficients, along with the application of a regularization technique, this method effectively addressed the inherent challenges of ill-posed inverse problems.

3) REINFORCEMENT LEARNING METHOD

As a progressive learning method, reinforcement learning (RL) is an interactive algorithm, in which a software agent with the ability to learn is rewarded for a set of actions when interacting with the environment. An environment comprised of a collection of potential states, denoted as S, and a set of actions, denoted as A, through which an agent can execute to modify its state. The agent engages with the environment by executing an action a_t at time step t while being in state s_t, and subsequently receiving a reward r_t. RL involves the process of determining the optimal actions to take in different states with the aim of maximizing the accumulated reward over a sequence of time steps, called the *return* G_t in (3):

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}, \tag{3}$$

where γ falls within the range (0, 1), and serves as a discount factor to avoid divergence of the infinite sum. The action-value function $Q_\pi(s, a)$ is also defined as

$$Q_\pi(s, \ a) = E_\pi[G_t|s_t = s, \ a_t = a], \tag{4}$$

which is the expected return assuming a policy π is followed.

Unlike other learning methods, the learner is not provided with explicit instructions on which actions to choose. Instead, it must explore and experiment with different actions to determine which ones result in the highest rewards. In complex scenarios, the consequences of an action may extend beyond immediate rewards and impact future situations, influencing all subsequent rewards. These two fundamental aspects, the trial-and-error nature of search and the delay in receiving rewards, are the key distinguishing characteristics of RL [53]. One can also refer to [54], which compares RL methods with EAs comprehensively.

Deep Q-network (DQN) is one of the most well-known RL methods using experience replay to improve the stability and efficiency of the learning process. By using a deep neural network (DNN) to estimate the Q-value function instead of the plain Q-value table, DQN enables the Q-learning algorithm to handle high-dimensional input problems. The Q-values are updated by approximating the Bellman equation as follows:

$$\begin{aligned} Q(s_t, a_t) \leftarrow\ & Q(s_t, a_t) \\ & + \alpha\Big[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\Big], \end{aligned} \tag{5}$$

where α represents the learning rate, and the other variables have been previously defined.

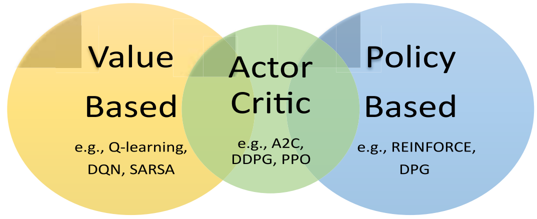


FIGURE 9. Classification of RL methods.

As depicted in Fig. 9, these methods belong to the family of value-based methods. In [55], Zhang et al. addressed the beam-steering issue of a phased-array antenna by developing a DQN-based system. The process involved training a DQN to determine the phase distribution of a conformal phased array antenna that corresponds to the desired beam steering angle. One significant benefit of employing deep RL (DRL)-based phased array antenna design is its ability to substantially decrease the time required to calibrate phase errors that naturally occur for each path in real-world situations. In another work [56], Wei et al. utilized a DQN framework to automatically design a decoupling metasurface for mutual coupling reduction between array antenna elements. The proposed method utilized the decision-making capacity of RL to acquire improved training data without requiring human intervention. However, one limitation of using the DQN algorithm is that it requires discretization of the action space, which can result in a higher computational complexity when addressing high-dimensional or continuous action space problems, and may lead to loss of important action information [53]. To mitigate this risk, more advanced RL algorithms such as deep deterministic policy gradient (DDPG) and proximal policy optimization (PPO) can be employed.

Policy-based RL methods are another family of RL methods directly focusing on learning a policy, which is a strategy or mapping from states to actions, without explicitly estimating a value function. The policy represents the agent's decision-making strategy in the environment. Policy-based methods, e.g., REINFORCE, deterministic policy gradient (DPG), etc., are particularly effective in scenarios with high-dimensional or continuous action spaces where traditional value-based methods might face challenges.

Actor-critic methods are a family of RL algorithms consisting of two major parts inherited from the value-based and policy-based methods: *actor* model and *critic* model. The actor model is updated using policy gradients that maximize the expected reward, while the critic model is updated using temporal difference (TD) learning that estimates the value of the current state. One of the benefits of the actor-critic algorithm is that it can learn from experience in real time, making it well-suited for online learning in dynamic environments. It is also computationally efficient so that it can handle high-dimensional state and action spaces. In [57], Peng et al. employed an actor-critic algorithm to

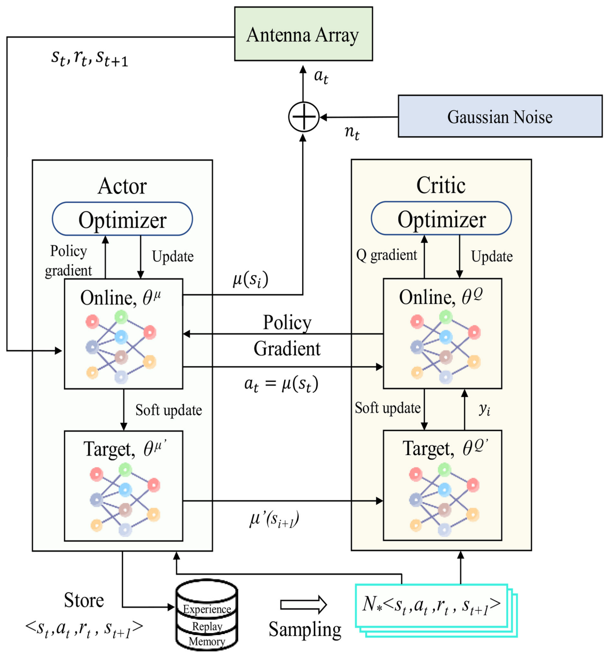
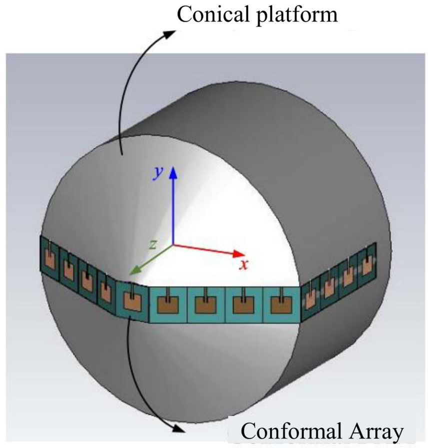


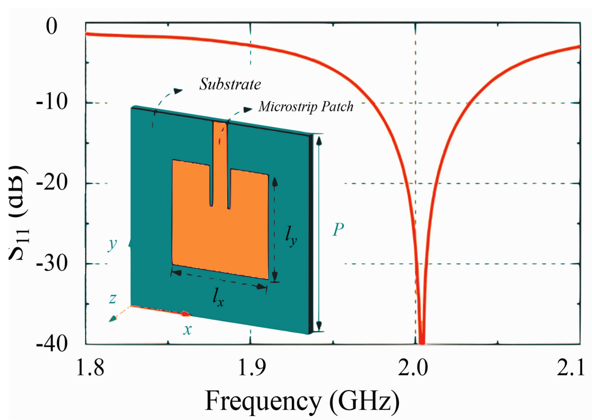
FIGURE 10. Procedure of the DDPG algorithm applied to design of the conformal array antenna adapted from [58].

TABLE 1. Optimization algorithms' performance [58].

Algorithm	GA	DDPG
Computational resources	CPU: 12% Memory: 1.27 GB	CPU: 27% Memory: 3.72 GB
Optimization time	11.92 hours per angle	98 ms per angle + 285 hours one-time training time
Sample size	N_p (population size)	Batch size 32



(a)



(b)

FIGURE 11. (a) Configuration of the conformal phased array antenna, (b) Reflection coefficient of the utilized patch antenna adapted from [58].

synthesize reflecting intelligent surfaces (RISs) to address the issue of radio frequency interference in the control area of a radio telescope. In another work, Zhang et al. [58] explored conformal phased array antenna (PAA) pattern synthesis across a broad angular range spanning from -150^{o} to 150° through the application of the DDPG algorithm as shown in Fig. 10. Here an actor and a critic interact with each other through the policy gradient mechanism. The configuration of the conformal PAA and the reflection coefficient of the utilized patch antenna are illustrated in Fig. 11(a) and (b), respectively. Table 1 shows a comparison between GA and the proposed DDPG algorithm with respect to computational resources, optimization time and sample size. One can easily conclude that DDPG is more successful in meeting the challenge of learning a complex problem, a process that demands significant time. Once it has mastered the complexity, DDPG can swiftly generate outputs for the next target beam steering-angle with high precision. Despite GA's computational and complexity advantages, it requires more training data along with significant recalculation and programming time for each beam scanning angle switch. Such limitation hinders its ability to provide quick and flexible radiation beam switching.

In a similar work, Zhang et al. [59] used the twin delayed deep deterministic policy gradient (TD3) algorithm to dramatically improve the stealth characteristic of a conformal array antenna. The outcomes from both simulation and measurement demonstrate remarkable adaptability in beam scanning, achieving a range of ±50 degree around 6 GHz. Additionally, it exhibits a noteworthy reduction of 10 dB in radar cross-section across the frequency bands of 3.1–5.3 GHz and 6.5–11.2 GHz.

C. HYBRID METHODS

Hybrid optimization methods combine different optimization techniques, typically integrating ML-based methods with non-ML-based methods, to create new and more powerful optimization methods. These hybrid methods leverage the strengths of different techniques to overcome the limitations of individual methods and improve overall optimization performance. Taking the prediction ability of ML methods in EAs, ML-assisted EAs are able to significantly reduce the effort spent on time-consuming EM simulations, when compared to the plain EA-based methods.

Surrogate-assisted EAs (SAEAs) are a type of optimization technique that combines surrogate modeling and EAs to efficiently solve complex optimization problems [60]. In [61], Liu et al. proposed a surrogate-model-assisted DE algorithm for antenna synthesis (SADEA). On top of its effectiveness over traditional algorithms, Akinsolu et al. [62] made further enhancements to evaluate multiple candidate designs by taking advantage of parallel simulation. This improvement resulted in 3.4-fold increase in speed compared to SADEA. In a similar work, Xue et al. [63] introduced a surrogate model-assisted hybrid optimization algorithm that employed the GP surrogate model-assisted DE and Nelder-Mead (NM) simplex algorithms.

It is common to adjust the geometry of microwave components to find optimal values for a given set of geometric parameters. This is called *geometry optimization*. In contrast, *topology optimization*, which aims to give more degrees of freedom to the optimizer to explore the optimal shape and topology of the EM structure, would help add or eliminate elements without any constraints on shape and topology. The conventional surrogate-based EM optimization approach is not suitable for solving the EM topology optimization problem due to its inability to relate different shapes and topologies by altering a fixed set of geometric parameters. Consequently, traditional surrogate models cannot adequately represent EM solutions when the design object comes to changes in topology. To overcome this limitation, Jin et al. [64] proposed an EM topology optimization technique tailored for microwave component design. They integrated Matrix-Padé-via-Lanczos with the Householder formula, simplifying the FEM matrix equation across frequencies to a single-frequency point for improving numerical efficiency. A novel technique leveraging GA's inheritance pattern further reduced the small matrix problem, exploiting shared characteristics between new and parent EM structures for additional computational cost reductions.

In [42], Fu et al. combined two ML-based approximation models with PSO to speed up prediction tasks. A simplified Kriging model was introduced to reduce the computational costs for high-dimensional datasets, while one RBF model was developed to predict new antenna structures. The experimental findings demonstrated that the search efficiency for the substrate integrated waveguide (SIW) and linear array antenna was enhanced by a factor of 2-4 when compared to the SADEA method. Therefore, selection of an appropriate optimization method is crucial for deriving optimal solutions in various domains. Some key performance criteria, including convergence rate, solution quality, and support for high-dimensional problems are used for the comparisons throughout Table 2.

TABLE 2. Comparison among different optimization method.

Method	Convergence Rate	Solution Quality	High-dimensional Problems
Genetic Algorithm [29]–[32]	slow	moderate	no
Particle Swarm Optimization [34]–[36]	moderate	moderate	no
Differential Evolution [37], [38]	moderate	moderate	no
Space Mapping [18]–[21]	fast	moderate	no
Homotopy Optimization [43], [44]	slow	high	no
Response Feature [22], [23]	moderate	high	no
Bayesian Optimization [24], [46]	moderate	moderate	yes
Support Vector Machine [47]	moderate	moderate	yes
K-means [48], [49]	slow	moderate	yes
Principal Component Analysis [50]–[52]	slow	high	yes
DQN [55], [56]	fast	high	yes
Actor-Critic methods [57]–[59]	fast	high	yes
SADEA [60]–[62]	moderate	moderate	no

IV. COMPONENT VIEW

Electromagnetic devices and circuits, such as antennas and filters, are essential components used in the modern era. In this section, we discuss optimization schemes for three of the most-common EM components: antennas, filters and metamaterials.

A. ANTENNAS

In the past decade, antennas have become crucial components for ensuring the optimal functionality of consumer electronics due to the growing popularity of mobile communication and Internet-of-Things (IoT). There is an increasing demand for antennas to be designed with upgraded precision, efficiency and complexity [47]. ML techniques have been successfully applied to the field of antennas to empower designers with greater CAD tools to facilitate the process of designing various antenna applications. For instance, in [65] a recurrent NN (RNN) model was utilized to generate appropriate complex weights for the antenna array beamformer to build a fast-tracking system. The generated weights from the RNN were then compared with the corresponding weights obtained through a null steering beamforming (NSB) technique to measure the RNN accuracy. The implementation of the RNN model demonstrated a high level of accuracy when adjusting the main lobe direction (i.e., main lobe divergence $< 0.5^\circ$) and positioning the nulls (i.e., nulls divergence $< 0.1^\circ$). This approach allows for the incorporation of some real-world factors, including the anisotropic radiation pattern of individual array elements and mutual coupling between elements.

In the next several years, there is a projected multi-fold increase in the commercial availability of 5G communication systems. These systems are driven by the utilization of mm-Wave active phased arrays (APAs) as the primary communication front end, serving as the key enabler for both 5G and the subsequent 6G evolution. The pursuit of maintaining quality of service (QoS) remains a significant driving force behind the advancement of communication systems. Consequently, the importance of fault diagnosis in communication systems has grown exponentially, as it plays a vital role in meeting user expectations regarding QoS [66]. As an example, in [67] Nielsen et al. used a DNN specifically designed to extract hidden features from the baseband in-phase and quadrature signals. Remarkably, it only requires a single probe at one measurement point for the diagnosis of faulty elements and components within APAs. The proposed method was validated using a commercial 28 GHz APA, demonstrating impressive accuracies of 99% for single-element and 80% for multi-element failure detections.

Direction-of-arrival (DOA) estimation, also known as *radio direction finding* (DF), finds diverse applications in various fields such as communications, remote sensing, and indoor localization, among others. This technique, widely used to determine the direction from which a signal is originated, has significant relevance in numerous domains. For instance, Friedrichs et al. utilized an ML-based method for high precision DOA estimation [68]. The effectiveness of the proposed method was showcased by applying it to an ultrawide-band circular array composed of miniaturized transverse electromagnetic (TEM) horns operating within a frequency range of 1.5 GHz to 5.5 GHz. In another work, Liu et al. in [69] introduced an autoencoder along with a DNN to preprocess the original array outputs for general DOA estimation. However, the proposed method tends to require a large amount of labeled data to train the DNN framework for DOA estimation. It may be very demanding in practical applications when such data is difficult to collect.

B. FILTERS

Among microwave circuits and modules, filters are critical because they possess distinct capabilities of transmitting and suppressing signals in order to function within particular frequency ranges. Filter design involves multiple steps and considerations, such as loss, bandwidth, operating frequency, stopband rejection, wideband performance, physical size, weight, operating power, and stability. Those must be tailored to meet the specific requirements for the practical applications.

A comprehensive survey of the state-of-the-art AI-assisted filter designs is provided in [10]. In [43], Zhao and Wu incorporated a homotopy algorithm along with an ANN for microwave and mmWave filter design. They provided two different five-pole rectangular waveguide design examples to meet the desired specifications at the center frequencies of 160GHz and 170GHz. However, this method may not be applicable to all types of filters including discrete, nonlinear and non-convex filters due to their discontinuous and non-smooth nature. In another work, Wu et al. proposed the use of an auxiliary neural network whose input includes not only electrical parameters but also partial physical/geometric variables [70]. Their result showed considerable error reduction within the same number of optimization iterations compared to the well-known DE optimization method.

In order to further facilitate the process of inverse modeling and address the issue of non-uniqueness, Zhang et al. in [71] presented a multivalued neural network, which enables the association of a single set of electrical parameters with multiple sets of geometric or physical parameters. This approach has the capability of effectively learning from training data by automatically considering contradictory information in the assignment of different values within the inverse model and eventually resolving the non-uniqueness problem. In [72], to handle the scenarios where the initial response for design optimization significantly deviates from the desired specifications, a multifeature-assisted neuro-TF was developed. The results showed more than 40 times faster convergence rate in comparison to direct EM optimization.

C. METAMATERIALS

Metamaterials, first introduced over 20 years ago, are a type of engineered EM material that can manipulate EM fields in unique ways. Over the past decade, there has been substantial growth and interest in EM metamaterials among physicists and engineers. These materials possess exceptional properties, including negative refraction, ultra-refraction, and anomalous dispersion, which do not naturally occur [73].

As a subset of metamaterials, 2-D structures of sub-wavelength unit cells made up of metallic scatterers and/or dielectric substrates are called *metasurfaces* [74]. The traditional approaches for designing metasurfaces heavily rely on the designer's intuition and previous experience. During the design process, selecting parameters in a multi-dimensional space is a time-consuming task and inevitably involves iterative loops through resource-intensive full-wave simulations. By virtue of their advantages, ANNs have been increasingly used as a viable option for metamaterial modeling. In [75], Yuan et al. proposed an improved TF-ANN model to expedite the inverse design of meta-surfaces. By using this method, they were able to reduce CPU time by over 4 times compared to a simple GA optimization, while achieving more than 20% higher accuracy than the direct inverse modeling method. Similar findings were reported in [76], where Koziel and Abdullah developed a rigorous ML-based framework followed by an expedited EM-driven fine-tuning process of metasurface geometric parameters. This approach resulted in a 15–25% enhancement in the bandwidth of radar cross-section (RCS) reduction compared to the original designs.

Similar to the inverse modeling methods in other EM areas, one approach to designing metasurfaces is inverse design. The inverse problem of predicting the physical composition of the metasurface structures based on the desired properties is not easy to solve due to the non-uniqueness challenge. For instance, in [77] Naseri and Hum proposed an encoder, a decoder and an ML model, which were trained to establish a generative ML-based approach for the inverse design of multilayer metasurfaces. However, one significant limitation of the metasurface inverse design is its considerable computational burden involved in the inverse design procedure. To address this issue, in [78] Szymanski et al. introduced a 2-D circuit network solver that utilized reduced-order models of the unit cells in a metastructure. This solver was employed alongside a gradient-based optimization algorithm, which utilized the adjoint variable method. This combined approach enabled the solution of large-scale optimization problems in the metastructure devices. In a similar report, Brown et al. in [79] aimed to tackle several key challenges involved in solving the inverse source problem. These include the formulation and optimization of a nonlinear cost function.

V. APPLICATION VIEW

In this section, we will explore three major application areas in microwave modeling; targeted by recent EM papers, namely; data generation, tolerance-aware design, and inverse modeling of EM components.

A. DATA GENERATION

To a large extent, the core challenge of direct optimization with full-wave EM simulations is the huge computational cost for determining the design variables and identifying the corresponding performance outputs. In the meantime, to train surrogate or other ML models, we need to generate the training data through extensive EM simulations [43]. Thus, the data generation stage seems to be inevitable. Moreover, as EM problems can be described by Maxwell's equations, much research has been devoted to solving Maxwell's equations by incorporating prior knowledge. The methods, such as FEM, MoM and FDM used in traditional solvers, rely on discretizing the space to solve equations. This means that there is a trade-off between resolution and speed. Calculations on coarse grids are faster but less accurate, whereas those on fine grids are accurate but slow. When dealing with complex partial differential equations (PDE) systems, a very fine discretization is often required, making it challenging and time-consuming for traditional solvers. In contrast, data-driven methods can learn trajectories of the equations directly from data, making them much faster than the conventional solvers, sometimes by orders of magnitude [80].

Machine learning methods have the potential to revolutionize the EM design automation and optimization domains by offering fast solvers that can approximate or improve upon traditional methods. However, classical neural networks are limited in that they can only learn solutions tied to a specific discretization scheme since they map between finite-dimensional spaces [81], [82], [83]. This could be a drawback for practical applications, and there is a need for development of mesh-invariant neural networks to overcome this limitation.

There are three principal neural-network based approaches for solving PDEs. The first approach uses finite-dimensional operators, each of which is parameterized as a deep CNN between finite-dimensional Euclidean spaces [84], [85], [86], [87]. This method intrinsically relies on meshes, which therefore requires adjustments and fine-tuning to achieve consistent errors across different resolutions and discretizations. On the other hand, the neural-FEM approach directly represents the solution function as a neural network. Its purpose is to model a single instance of PDE, rather than the solution operators [88], [89], [90]. This method is not reliant on meshes, which means it is accurate and independent of mesh-related issues. However, if a new instance of functional parameter or coefficient is encountered, a new neural network must be trained.

In time-domain computational EM, the propagation of EM waves can be visualized as a sequence of color images, akin to a video. Each cross-section of the simulated space consists of three field components represented as pixel grids forming monochrome images. These images combine into a color image that evolves over time, creating a video sequence. This method connects computational EM with computer vision, enabling rapid and efficient predictions of EM wave propagation in new configurations based on learned patterns. For instance, Noakoasteen et al. [91] employed an encoder-recurrent-decoder architecture trained on finite-difference time-domain (FDTD) simulations of plane wave scattering from distributed, perfect electric conductor (PEC) objects. Their demonstrations showed that the trained network significantly accelerated simulation time, achieving more than 17 times speed-up compared to traditional FDTD solvers for transient electrodynamics problems. As explained in Section I, the transformer architectures excel in time-series optimization by adeptly modeling complex relationships and capturing long-range dependencies. Their efficient parallel processing capabilities significantly enhance the accuracy and speed of EM design and optimization tasks. In [92], a transformer model and a Convolutional Graph Neural Network (CGNN) were used to emulate the dynamics of EM fields propagating and scattering from PEC objects, achieving computation speed-ups of 14x and 9x, respectively.

In the transformer architectures as illustrated in Fig. 12, the encoder processes input sequences, converting them into a series of embeddings that represent each token's semantics and position. These embeddings are enhanced with positional encoding to preserve sequence order. The decoder, on the other hand, generates output sequences based on the processed input and context, utilizing attention mechanisms to focus on relevant input tokens. Embeddings are vector representations of tokens, capturing their semantic meaning within the sequence. At the output layer, a softmax function generates probability distributions over the

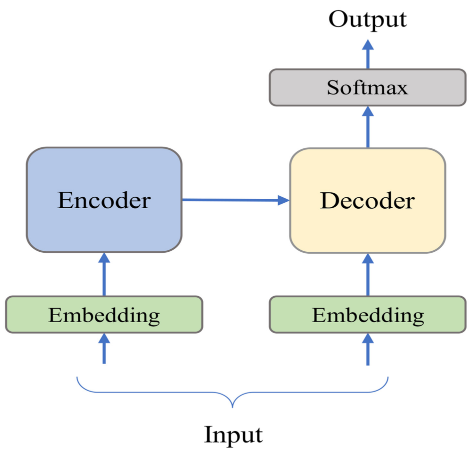


FIGURE 12. Architecture of a transformer model.

vocabulary, facilitating the model's generation of coherent and meaningful output sequences by assigning probabilities to each token in the vocabulary. This is particularly useful for modeling nonlinear relationships in EM data, where interactions between design variables can significantly impact the EM performance.

Fotiadis et al. [93] utilized a U-Net architecture to approximate wave propagation, reducing the long-term approximation RMSE to 0.071 from a previous baseline of 0.186 and achieving a 240-fold increase in speed over traditional simulations. In a related study [94], Mohan et al. introduced the Compressed Convolutional LSTM (CC-LSTM) framework, which combined convolutional autoencoders and LSTM networks for efficient, high-fidelity modeling of three-dimensional turbulent flows, resulting in significant computational savings. Additionally, Sorteberg et al. [95] presented a neural network model designed to predict wave propagation in a two-dimensional medium, demonstrating accurate predictions up to 80 frames from limited initial observations. This model, which integrates an encoder, LSTM propagators, and a decoder, generalizes well to new initial conditions and significantly accelerates simulation time. While the numerical algorithm takes an average of 35 seconds, the network can generate predictions for 80 frames in less than half a second.

A new approach, called *neural operators*, has emerged by using neural networks to learn mesh-free and infinite-dimensional operators [96], [97], [98], [99]. This new technique overcomes the mesh-dependent nature of the finite-dimensional operator methods described above by generating a single set of network parameters that can be used with different discretizations. This neural operator, which can transfer solutions between meshes, only needs to be trained once. To obtain a solution for a new instance of parameters, only a forward pass of the network is required, which avoids the computational issues experienced in the neural-FEM methods. It has been recently shown that Fourier neural operators (FNOs) have great potential to accelerate the solving of PDEs. In fact, neural operators can directly learn how to map any parametric dependence of a function to its solution. That is to say, they can learn a whole range of PDEs, unlike traditional methods which only solve a single instance of equations [100], [101]. For instance, Zhang and Rahmat-Samii in [102] employed a neural network equipped with FNOs to predict the electric potential in 2D models with varying permittivity distributions. It demonstrated exceptional precision, outperforming the traditional FDM by over 1,000 times in runtime. This showcases the significant promise of neural network operators in solving the time-variant Maxwell's equations.

B. TOLERANCE-AWARE DESIGN/SYNTHESIS

In real-world EM component design and synthesis, one needs to take some uncertainties into consideration to avoid failing to meet the desired specifications. These uncertainties may stem from either fabrication tolerances or lack of precise knowledge of material parameters. As a result, reliable uncertainty quantification (UQ) is important when EM simulation tools are employed to evaluate full-wave EM simulation models, which impose great computational costs. It becomes even more vital when deviations of geometry parameters from their nominal values lead to frequency shifts in the operating frequency band, especially in microwave and mmWave component design [22], [103], [104], [105], [106], [107], [108].

In the realm of EM design, the traditional approaches of parameter sweeping, or trial-and-error based on domain knowledge, can be very time-consuming, normally offering no assurance of finding an optimal solution. Examining the complex relationships between input and output variables, particularly in the scenarios with high-dimensional design and performance spaces, proves highly challenging even when using high-fidelity models. To reduce the exploration time in the design space, one strategy is to use the surrogate models (as discussed in Section II-A), which are low-fidelity approximations of a high-fidelity model [100]. Notwithstanding that, the quality of the surrogate model has a crucial impact on the convergence rate as well as the computational cost.

In order to avoid poor-quality surrogate modeling and take advantage of each individual surrogate model, an ensemble of surrogate models can be constructed with a weight coefficient given to each model to form a linear combination of surrogate models [14]. To further improve the balance between convergence rate and computational complexity, Chen et al. [109] proposed to use a multibranch ML-assisted optimization (MB-MLAO) scheme to significantly decrease the computational complexity involved in the EM optimization task. Then the approach was applied to the antenna design searching for worst-case performance (WCP) with a consideration of realistic manufacturing tolerances. However, when dealing with a high number of parameters, typically in the range of several dozens, the curse of dimensionality becomes a significant limiting factor.

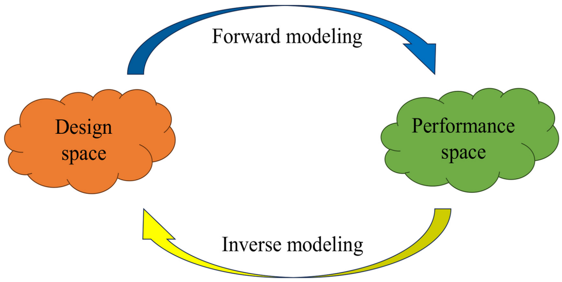


FIGURE 13. Forward modeling versus inverse modeling.

In [110], Zhang et al. combined quadratic mappings and matrix-valued transfer functions to create a surrogate model that accelerated the estimation and optimization of yield for multiport microwave structures. The proposed method was illustrated by using three microwave structures operating at a center frequency of 12GHz, showcasing the benefits and advantages of the approach.

C. INVERSE MODELING

An inverse problem refers to the process of extracting or designing physical or geometric parameters by analyzing the desired outputs such as gain, bandwidth and S-parameters in a reverse manner, with inputs and outputs switched. Forward modeling involves training a model such as a neural network model to simulate a microwave component. Fig. 13 provides a visual illustration of both the forward and inverse modeling processes. In the forward modeling, the inputs are physical or geometric parameters, while the outputs are electrical parameters.

The forward model can be mathematically expressed as a function:

$$y = f(x), \tag{6}$$

where x and y represent the input and output of the forward model, respectively, and f stands for the function representing the forward model. However, in the inverse modeling, the formulation of the relationship between inputs and outputs can be determined by:

$$x = f^{-1}(y) = g(y), \tag{7}$$

where g represents the inverse model function. Once a reliable inverse model has been developed, it allows for the direct extraction of the physical or geometric parameters without a need for iterative optimization processes. As a result, the inverse modeling offers a faster and more efficient approach compared to traditional iterative optimization methods [71].

Nevertheless, when it comes to common scenarios, there is often a lack of uniqueness in the inverse relationship between input and output in the training data. This is why training a model to predict inputs (design variables) from outputs (electrical performances) is generally more difficult than training a model to predict outputs from inputs. To address the non-uniqueness problem of antenna inverse modeling, Xiao et al. in [111] proposed a novel structure for the ANN model. Their method showed over 10 times faster convergence rate in practical examples such as multimode resonant antenna and triband printed antenna with twelve and twenty-four design variables, respectively.

TABLE 3. Number of publications on different combinations of EM components and different ML-based methods since 2013 sourced from IEEE Xplore, ACM, Elsevier ScienceDirect, and Springer databases.

		Antennas	Filters	Metamaterials
ML methods (excluding RL)	Bayesian optimization	15	1	0
	Neural networks	191	148	19
	SVMs	26	25	2
	Gaussian process	47	25	0
RL methods	Mostly DQN	20	5	1

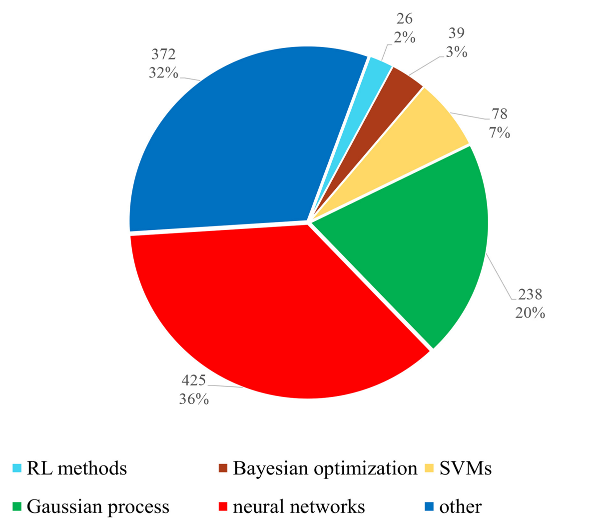


FIGURE 14. Total numbers and percentages of publications within EM area for the ML methods over the past decade sourced from IEEE Xplore, ACM, Elsevier ScienceDirect, and Springer databases.

VI. STATISTICS AND PERSPECTIVES

As we delve into a comprehensive exploration aimed at uncovering the most promising strategies for addressing EM optimization and design challenges, we have combined the terminologies of different ML methods versus the RL methods with three chosen EM components (including antennas, filters and metamaterials), as listed in Table 3. It turns out that there are few papers that utilize the RL methods for designing antennas, filters, and metamaterials, while using other ML methods seems to be more common among researchers.

The accumulated numbers of publications in the last decade within EM area on the topics of RL, Bayesian optimization, SVMs, Gaussian process, and neural networks are analyzed and compared across IEEE Xplore, ACM, Elsevier ScienceDirect, and Springer databases in Fig. 14. It

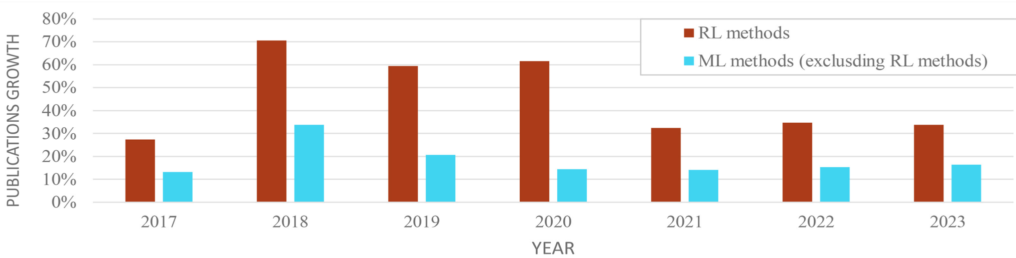


FIGURE 15. Publication growth on various ML-based methods (excluding RL) versus RL methods during the recent seven years sourced from IEEE Xplore, ACM, Elsevier ScienceDirect, and Springer databases.

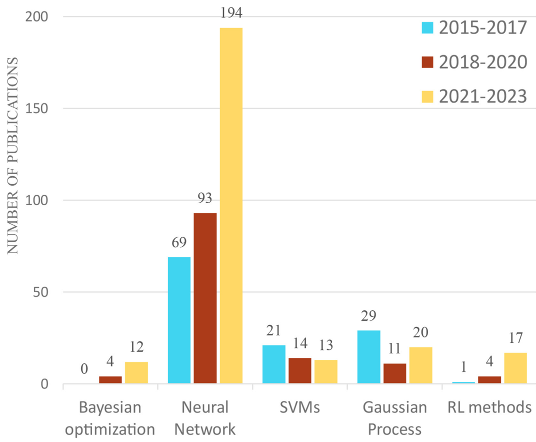


FIGURE 16. Publications on various ML-based methods in the EM area consisting of antennas, filters, and metamaterials, across different time periods sourced from IEEE Xplore, ACM, Elsevier ScienceDirect, and Springer databases.

can be observed that the total number of publications on RL methods only constitutes nearly 2% of the total number of publications in the field of EM design. In contrast, Bayesian methods hold around 3%, while SVMs, Gaussian process and neural networks contribute 7%, 20%, and 36%, respectively. Additionally, other ML methods account for approximately one third of the total publications.

To gain deeper insights into the evolving landscape of research trends, Fig. 15 depicts the growth of publications in ML (excluding RL) methods compared to RL methods within the field of Engineering and Computer Science (ECS) over the past seven years. The results highlight a substantial increase in the attention garnered by RL methods as a prominent technique within the ECS research communities.

In recent years, there has been a notable trend of integrating ML techniques into the EM design process, especially with RL, demonstrating significant potential for enhancing optimization and automation performance within the EM area. We have also investigated publications on different ML methods employed in the EM area during various time periods, as illustrated in Fig. 16. One can observe that the numbers of publications for EM component design on Bayesian optimization, neural networks, and RL methods are increasing, while those for other ML methods are fairly stable. The statistical data also underline the fact that integration of the RL techniques into the optimization of such EM components is currently undergoing an initial phase of research and development, compared to other ML methods.

Over the past few years, EM structures such as antennas and filters have continued to gain significant attention and remained popular within the engineering community. Interestingly, there has been a noticeable surge in the study towards metamaterial design in recent years, representing a noteworthy shift in the EM research landscape. The growing popularity of metamaterial design is driven by its remarkable ability to manipulate electromagnetic waves with unprecedented precision and efficiency, especially applicable to beamforming, antenna miniaturization, polarization manipulation and frequency filtering.

There are tremendous opportunities for further innovations and applications of ML methods for EM modeling and design, from advanced RL algorithms to novel data generating methods. Further advances in quantum RL methods emerge to be another important direction, opening up a new window to the ML area, as it has shown its capabilities in solving other complex problems [112], [113], [114], [115], [116]. Moreover, the process of design is naturally an inverse problem of performance analysis. Therefore, the pursuit of inverse modeling for straightforward design solutions remains an appealing avenue of research. Incorporation of multi-agent RL methods and game theory into solutions of high-dimensional multi-objective EM problems may lead to new approaches to conquer the dilemma of the EM optimization challenges.

VII. CONCLUSION

In this paper, a comprehensive review of recent optimization techniques in EM modeling was conducted. Since the early work on space mapping and population-based algorithms, this field of research has drawn plenty of interest in both industry and academia. On top of surveying from structural, algorithm, application and component views, we delved into the literature statistics and explored future prospects. The current progress of research and development demonstrated that ML techniques, especially RL, are promising, mainly by virtue of their strong capability to explore high-dimensional input-variable spaces and multi-objective output performance spaces. Looking ahead, we deemed quantum computation to become instrumental in handling high computational tasks, thereby facilitating the design and optimization processes for solving EM challenges.

REFERENCES

[1] J. E. Rayas-Sanchez, S. Koziel, and J. W. Bandler, "Advanced RF and microwave design optimization: A journey and a vision of future trends," *IEEE J. Microw.*, vol. 1, no. 1, pp. 481–493, Jan. 2021.

[2] S. Nabavi and L. Zhang, "Design and optimization of MEMS piezoelectric energy harvesters for improved efficiency," in *Proc. IEEE Sens. Conf.*, 2017, pp. 1–3.

[3] K. Sharma and G. P. Pandey, "Efficient modelling of compact microstrip antenna using machine learning," *AEU Int. J. Electron. Commun.*, vol. 135, Jun. 2021, Art. no. 153739.

[4] Q. Xu, S. Zeng, F. Zhao, R. Jiao, and C. Li, "On formulating and designing antenna arrays by evolutionary algorithms," *IEEE Trans. Antennas Propag.*, vol. 69, no. 2, pp. 1118–1129, Feb. 2021.

[5] Z. Zhao and L. Zhang, "Deep reinforcement learning for analog circuit structure synthesis," in *Proc. IEEE Design, Autom. Test Europe Conf. Exhibition*, 2022, pp. 1–4.

[6] T. Liao and L. Zhang, "Efficient parasitic-aware gm/ID-based hybrid sizing methodology for analog and RF integrated circuits," *ACM Trans. Design Autom. Electron. Syst.*, vol. 26, no. 2, pp. 1–31, 2020.

[7] H. M. El Misilmani, T. Naous, and S. K. Al Khatib, "A review on the design and optimization of antennas using machine learning algorithms and techniques," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 30, no. 10, pp. 1–28, Jul. 2020.

[8] Q. Wu, Y. Cao, H. Wang, and W. Hong, "Machine-learning-assisted optimization and its application to antenna designs: Opportunities and challenges," *China Commun.*, vol. 17, no. 4, pp. 152–164, Apr. 2020.

[9] S. D. Campbell, R. P. Jenkins, P. J. O'Connor, and D. Werner, "The explosion of artificial intelligence in antennas and propagation: How deep learning is advancing our state of the art," *IEEE Antennas Propag. Mag.*, vol. 63, no. 3, pp. 16–27, Jun. 2021.

[10] Y. Yu et al., "State-of-the-art: AI-assisted surrogate modeling and optimization for microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4635–4651, Nov. 2022.

[11] F. Feng, W. Na, J. Jin, J. Zhang, W. Zhang, and Q. J. Zhang, "Artificial neural networks for microwave computer-aided design: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4597–4619, Nov. 2022.

[12] A. Seretis and C. D. Sarris, "An overview of machine learning techniques for radiowave propagation modeling," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3970–3985, Jun. 2022.

[13] Y. O. Cha, A. A. Ihalage, and Y. Hao, "Antennas and propagation research from large-scale unstructured data with machine learning: A review and predictions," *IEEE Antennas Propag. Mag.*, vol. 65, no. 5, pp. 10–24, Oct. 2023.

[14] P. Jiang, Q. Zhou, and X. Shao, *Surrogate Model-Based Engineering Design and Optimization*. Singapore: Springer, 2020.

[15] B. Liu et al., "An efficient method for complex antenna design based on a self adaptive surrogate model-assisted optimization technique," *IEEE Trans. Antennas Propag.*, vol. 69, no. 4, pp. 2302–2315, Apr. 2021.

[16] Y. Liu et al., "An efficient method for antenna design based on a self-adaptive Bayesian neural network assisted global optimization technique," *IEEE Trans. Antennas Propag.*, vol. 70, no. 12, pp. 11375–11388, Dec. 2022.

[17] R. Shibata, M. Ohira, and Z. Ma, "A novel convolutional-autoencoder based surrogate model for fast S-parameter calculation of planar BPFs," in *IEEE/MTT-S Int. Microw. Symp. Dig.*, Jun 2022, pp. 498–501.

[18] J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 42, no. 12, pp. 2536–2544, Dec. 1994.

[19] W. Zhang et al., "Advanced parallel space-mapping-based multiphysics optimization for high-power microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 5, pp. 2470–2484, May 2021.

[20] J. C. Melgarejo, J. Ossorio, S. Cogollos, M. Guglielmi, V. E. Boria, and J. W. Bandler, "On space mapping techniques for microwave filter tuning," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 12, pp. 4860–4870, Dec. 2019.

[21] W. Zhang et al., "Space mapping approach to electromagnetic centric multiphysics parametric modeling of microwave components," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3169–3185, Jul. 2018.

[22] S. Koziel and A. Pietrenko-Dabrowska, "Tolerance optimization of antenna structures by means of response feature surrogates," *IEEE Trans. Antennas Propag.*, vol. 70, no. 11, pp. 10988–10997, Nov. 2022.

[23] A. Pietrenko-Dabrowska and S. Koziel, "Generalized formulation of response features for reliable optimization of antenna input characteristics," *IEEE Trans. Antennas Propag.*, vol. 70, no. 5, pp. 3733–3748, May 2022.

[24] J. Zhou et al., "A trust-region parallel Bayesian optimization method for simulation-driven antenna design," *IEEE Trans. Antennas Propag.*, vol. 69, no. 7, pp. 3966–3981, Jul. 2021.

[25] Y. Zhuo, F. Feng, J. Zhang, and Q. J. Zhang, "Parametric modeling incorporating joint polynomial-transfer function with neural networks for microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4652–4665, Nov. 2022.

[26] R. Garnett, *Bayesian Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2023.

[27] Z. Zhao and L. Zhang, "Efficient performance modeling for automated CMOS analog circuit synthesis," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 29, no. 11, pp. 1824–1837, Nov. 2021.

[28] M. Baranowski, G. Fotyga, A. Lamecki, and M. Mrozowski, "Bayesian optimization for solving high-frequency passive component design problems," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 70, no. 4, Aug. 2022, Art. no. e141595.

[29] R. M. Bichara, F. A. Z. Asadallah, M. Awad, and J. Costantine, "A miniaturized reconfigurable antenna using quantum genetic algorithm optimization," in *Proc. IEEE Int. Symp. Antennas Propag. North Amer. Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 763–764.

[30] T. Akada and K. Fujimori, "Designing microwave circuits using genetic algorithms accelerated by convolutional neural networks," in *Proc. 50th Eur. Microw. Conf. (EuMC)*, Jan. 2021, pp. 61–64.

[31] L. Zhang, R. Raut, L. Wang, and Y. Jiang, "Analog module placement realizing symmetry constraints based on a radiation decoder," in *Proc. IEEE Int. Midwest Symp. Circuits Syst.*, 2004, pp. I481–I484.

[32] J. S. Smith and M. E. Baginski, "Thin-wire antenna design using a novel branching scheme and genetic algorithm optimization," *IEEE Trans. Antennas Propag.*, vol. 67, no. 5, pp. 2934–2941, May 2019.

[33] Q. Yu, Y. Guan, Y. Yu, and C. Yu, "Genetic algorithm optimized back propagation neural networks in behavioral modeling of power amplifiers excited by 5G NR signal," in *Proc. IEEE Asia–Pacific Conf. Antennas Propag. (APCAP)*, Aug. 2020, pp. 1–2.

[34] R. Cui, Z. Mei, and T. Niu, "An improved particle swarm optimization for antenna design," in *Proc. IEEE Int. Symp. Antennas Propag. North Amer. Radio Sci. Meeting (APS/URSI)*, 2021, pp. 1457–1458.

[35] X. Luo, B. Yang, and H. J. Qian, "Adaptive synthesis for resonator-coupled filters based on particle swarm optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 2, pp. 712–725, Feb. 2019.

[36] C. X. Yang, J. Zhang, and M. S. Tong, "A hybrid quantum-behaved particle swarm optimization algorithm for solving inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 69, no. 9, pp. 5861–5869, Sep. 2021.

[37] B. Liu, H. Yang, and M. J. Lancaster, "Synthesis of coupling matrix for diplexers based on a self-adaptive differential evolution algorithm," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 2, pp. 813–821, Feb. 2018.

[38] C. Y. Cui, Y. C. Jiao, L. Zhang, L. Lu, and H. Zhang, "Synthesis of subarrayed monopluse arrays with contiguous elements using a DE algorithm," *IEEE Trans. Antennas Propag.*, vol. 65, no. 8, pp. 4340–4345, Aug. 2017.

[39] X. Li and K. M. Luk, "The grey wolf optimizer and its applications in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 2186–2197, Mar. 2020.

[40] A. D. Boursianis, S. K. Goudos, M. Matthaiou, S. Koulouridis, and M. Salucci, "Implantable antenna design using improved grey wolf optimizer algorithm," in *Proc. IEEE Int. Symp. Antennas Propag. North Amer. Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 801–802.

[41] X. Li and Y. X. Guo, "Multiobjective optimization design of aperture illuminations for microwave power transmission via multiobjective grey wolf optimizer," *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 6265–6276, Aug. 2020.

[42] K. Fu, X. Cai, B. Yuan, Y. Yang, and X. Yao, "An efficient surrogate assisted particle swarm optimization for antenna synthesis," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 4977–4984, Jul. 2022.

[43] P. Zhao and K. Wu, "Homotopy optimization of microwave and millimeter-wave filters based on neural network model," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 4, pp. 1390–1400, Apr. 2020.

[44] C. Roy and K. Wu, "Homotopy optimization and ANN modeling of millimeter-wave SIW cruciform coupler," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4751–4764, Nov. 2022.

[45] A. Abouzarkhanifard, H. Chimeh, M. Al Janaideh, T. Zou, and L. Zhang, "Transfer-learning-aided optimization for a low-frequency wideband MEMS energy harvester," in *Proc. IEEE Sens. Conf.*, 2022, pp. 1–4.

[46] T. Liao and L. Zhang, "High-dimensional many-objective Bayesian optimization for LDE-aware analog IC sizing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 30, no. 1, pp. 15–28, Jan. 2022.

[47] D. Shi, C. Lian, K. Cui, Y. Chen, and X. Liu, "An intelligent antenna synthesis method based on machine learning," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 4965–4976, Jul. 2022.

[48] P. Lyu, A. Benlarbi-Delai, Z. Ren, and J. Sarrazin, "Angular clustering of millimeter-wave propagation channels with watershed transformation," *IEEE Trans. Antennas Propag.*, vol. 70, no. 2, pp. 1279–1290, Feb. 2022.

[49] Z. Zhang, H. Chen, F. Jiang, Y. Yu, and Q. S. Cheng, "K-means-based multigroup differential evolution optimization framework for design of MIMO antenna with decoupling elements," *IEEE Antennas Wireless Propag. Lett.*, vol. 21, no. 10, pp. 1980–1984, Oct. 2022.

[50] I. T. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer, 2002.

[51] R. Reris and J. P. Brooks, "Principal component analysis and optimization: A tutorial," in *Proc. 14th INFORMS Comput. Soc. Conf.*, Feb. 2015, pp. 212–225.

[52] M. Sedaghat, R. Trinchero, and F. Canavero, "Compressed machine learning-based inverse model for the design of microwave filters," in *IEEE/MTT-S Int. Microw. Symp. Dig.*, Jun. 2021, pp. 13–15.

[53] M. Ahmadi and L. Zhang, "Analog layout placement for FinFET technology using reinforcement learning," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2021, pp. 1–5.

[54] A. Y. Majid, S. Saaybi, V. Francois-Lavet, R. V. Prasad, and C. Verhoeven, "Deep reinforcement learning versus evolution strategies: A comparative survey," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 2, 2023, doi: 10.1109/TNNLS.2023.3264540.

[55] B. Zhang, C. Jin, K. Cao, Q. Lv, and R. Mittra, "Cognitive conformal antenna array exploiting deep reinforcement learning method," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5094–5104, Jul. 2022.

[56] Z. Wei et al., "Fully automated design method based on reinforcement learning and surrogate modelling for antenna array decoupling," *IEEE Trans. Antennas Propag.*, vol. 71, no. 1, pp. 660–671, Jan. 2022.

[57] J. Peng, H. Cao, Z. Ali, X. Wu, and J. Fan, "Intelligent reflecting surface-assisted interference mitigation with deep reinforcement learning for radio astronomy," *IEEE Antennas Wireless Propag. Lett.*, vol. 21, no. 9, pp. 1757–1761, Sep. 2022.

[58] B. Zhang et al., "Ultra-wide-scanning conformal heterogeneous phased array antenna based on deep deterministic policy gradient algorithm," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5066–5077, Jul. 2022.

[59] B. Zhang, C. Jin, B. Tian, K. Cao, J. Wang, and P. Zhang, "Electromagnetic-model-driven twin delayed deep deterministic policy gradient algorithm for stealthy conformal array antenna," *IEEE Trans. Antennas Propag.*, vol. 70, no. 12, pp. 11779–11789, Oct. 2022.

[60] B. Liu, H. Aliakbarian, Z. Ma, G. A. E. Vandenbosch, G. Gielen, and P. Excell, "An efficient method for antenna design optimization based on evolutionary computation and machine learning techniques," *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 7–18, Jan. 2014.

[61] B. Liu, M. O. Akinsolu, N. Ali, and R. Abd-Alhameed, "Efficient global optimization of microwave antennas based on a parallel surrogate model-assisted evolutionary algorithm," *IET Microw., Antennas Propag.*, vol. 13, no. 2, pp. 149–155, Feb. 2019.

[62] M. O. Akinsolu, B. Liu, V. Grout, P. I. Lazaridis, M. E. Mognaschi, and P. Di Barba, "A parallel surrogate model assisted evolutionary algorithm for electromagnetic design optimization," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 3, no. 2, pp. 93–105, Apr. 2019.

[63] L. Xue, B. Liu, Y. Yu, Q. S. Cheng, M. Imran, and T. Qiao, "An unsupervised microwave filter design optimization method based on a hybrid surrogate model-assisted evolutionary algorithm," *IEEE Trans. Microw. Theory Techn.*, vol. 71, no. 3, pp. 1159–1170, Mar. 2022.

[64] J. Jin, F. Feng, J. Zhang, J. Ma, and Q. J. Zhang, "Efficient EM topology optimization incorporating advanced matrix Padé via Lanczos and genetic algorithm for microwave design," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 8, pp. 3645–3666, Aug. 2021.

[65] I. Mallioras, Z. D. Zaharis, P. I. Lazaridis, and S. Pantelopoulos, "A novel realistic approach of adaptive beamforming based on deep neural networks," *IEEE Trans. Antennas Propag.*, vol. 70, no. 10, pp. 8833–8848, Oct. 2022.

[66] K. P. Prajosh, S. S. Ranganathan, F. Ferranti, and U. K. Khankhoje, "Efficient mutual-coupling aware fault diagnosis of phased array antennas using optimized excitations," *IEEE Antennas Wireless Propag. Lett.*, vol. 21, no. 9, pp. 1906–1910, Sep. 2022.

[67] M. H. Nielsen et al., "Robust and efficient fault diagnosis of mm-Wave active phased arrays using baseband signal," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5044–5053, Jul. 2022.

[68] G. R. Friedrichs, M. A. Elmansouri, and D. S. Filipovic, "A compact machine learning architecture for wideband amplitude-only direction finding," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5189–5198, Jul. 2022.

[69] Z. M. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.

[70] Y. Wu, G. Pan, D. Lu, and M. Yu, "Artificial neural network for dimensionality reduction and its application to microwave filters inverse modeling," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4683–4693, Nov. 2022.

[71] C. Zhang, J. Jin, W. Na, Q. J. Zhang, and M. Yu, "Multivalued neural network inverse modeling and applications to microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 8, pp. 3781–3797, Aug. 2018.

[72] F. Feng et al., "Multifeature-assisted neuro-transfer function surrogate-based EM optimization exploiting trust-region algorithms for microwave filter design," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 2, pp. 531–542, Feb. 2020.

[73] Z. Fang and J. Zhan, "Deep physical informed neural networks for metamaterial design," *IEEE Access*, vol. 8, pp. 24506–24513, 2020.

[74] Z. Wei et al., "Equivalent circuit theory-assisted deep learning for accelerated generative design of metasurfaces," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5120–5129, Jul. 2022.

[75] L. Yuan et al., "An efficient artificial neural network model for inverse design of metasurfaces," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, no. 6, pp. 1013–1017, Jun. 2021.

[76] S. Koziel and M. Abdullah, "Machine-learning-powered EM-based framework for efficient and reliable design of low scattering metasurfaces," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 4, pp. 2028–2041, Apr. 2021.

[77] P. Naseri and S. V. Hum, "A generative machine learning-based approach for inverse design of multilayer metasurfaces," *IEEE Trans. Antennas Propag.*, vol. 69, no. 9, pp. 5725–5739, Sep. 2021.

[78] L. Szymanski, G. Gok, and A. Grbic, "Inverse design of multi-input multi-output 2-D metastructured devices," *IEEE Trans. Antennas Propag.*, vol. 70, no. 5, pp. 3495–3505, May 2022.

[79] T. Brown, C. Narendra, Y. Vahabzadeh, C. Caloz, and P. Mojabi, "On the use of electromagnetic inversion for metasurface design," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 1812–1824, Mar. 2020.

[80] Z. Li et al., "Fourier neural operator for parametric partial differential equations," in *Proc. Int. Conf. Learn. Represent.*, Jan. 2021, pp. 1–16.

[81] C. M. Jiang et al., "MeshfreeFlowNet: A physics-constrained deep continuous space-time super-resolution framework," in *Proc. Int. Conf. High Perform. Comput., Netw., Storage Anal.*, Nov. 2020, pp. 1–16.

[82] D. Greenfeld, M. Galun, R. Kimmel, I. Yavneh, and R. Basri, "Learning to optimize multigrid PDE solvers," in *Proc. the 36th Int. Conf. Mach. Learn.*, Jun. 2019, pp. 2415–2423.

[83] D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer, "Machine learning accelerated computational fluid dynamics," *Proc. Nat. Acad. Sci.*, vol. 118, no. 21, May 2021, Art. no. e2101784118.
[84] Y. Khoo, J. Lu, and L. Ying, "Solving parametric PDE problems with artificial neural networks," *Eur. J. Appl. Math.*, vol. 32, no. 3, pp. 421–435, Jul. 2021.
[85] S. Bhatnagar, Y. Afshar, S. Pan, K. Duraisamy, and S. Kaushik, "Prediction of aerodynamic flow fields using convolutional neural networks," *Comput. Mech.*, vol. 64, no. 2, pp. 525–545, Aug. 2019.
[86] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Probl. J.*, vol. 32, no. 12, Apr. 2017, Art. no. 124007.
[87] X. Guo, W. Li, and F. Iorio, "Convolutional neural networks for steady flow approximation," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, Aug. 2016, pp. 481–490.
[88] E. Weinan and B. Yu, "The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems," *Commun. Math. Stat.*, vol. 6, pp. 1–12, Feb. 2018.
[89] L. Bar and N. Sochen, "Unsupervised deep learning algorithm for PDE-based forward and inverse problems," Apr. 2019, *arXiv:1904.05417*.
[90] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, Feb. 2019.
[91] O. Noakoasteen, S. Wang, Z. Peng, and C. Christodoulou, "Physics-informed deep neural networks for transient electromagnetic analysis," *IEEE Open J. Antennas Propag.*, vol. 1, no. 1, pp. 404–412, Aug. 2020.
[92] O. Noakoasteen, C. Christodoulou, Z. Peng, and S. K. Goudos, "Physics-informed surrogates for electromagnetic dynamics using transformers and graph neural networks," *IET Microw., Antennas Propag.*, to be published.
[93] S. Fotiadis, E. Pignatelli, M. L. Valencia, C. Cantwell, A. Storkey, and A. A. Bharath, "Comparing recurrent and convolutional neural networks for predicting wave propagation," Feb. 2020, *arXiv:2002.08981*.
[94] A. Mohan, D. Daniel, M. Chertkov, and D. Livescu, "Compressed convolutional LSTM: An efficient deep learning framework to model high fidelity 3D turbulence," Feb. 2019, *arXiv:1903.00033*.
[95] W. E. Sorteberg, S. Garasto, A. S. Pouplin, C. D. Cantwell, and A. A. Bharath, "Approximating the solution to wave propagation using deep neural networks," Dec. 2018, *arXiv:1812.01609*.
[96] R. G. Patel, N. A. Trask, M. A. Wood, and E. C. Cyr, "A physics-informed operator regression framework for extracting data-driven continuum models," *Comput. Methods Appl. Mech. Eng.*, vol. 373, Jan. 2021, Art. no. 113500.
[97] Z. Li et al., "Multipole graph neural operator for parametric partial differential equations," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.*, Dec. 2020, pp. 6755–6766.
[98] Z. Li et al., "Neural Operator: Graph kernel network for partial differential equations," Mar. 2020, *arXiv:2003.03485*.
[99] K. Bhattacharya, B. Hosseini, N. B. Kovachki, and A. M. Stuart, "Model reduction and neural networks for parametric PDEs," *SMAI J. Comput. Math.*, vol. 7, pp. 121–157, May 2020.
[100] Z. Li, D. Z. Huang, B. Liu, and A. Anandkumar, "Fourier neural operator with learned deformations for PDEs on general geometries," Jul. 2022, *arXiv:2207.05209*.
[101] C. Zhu, H. Ye, and B. Zhan, "Fast solver of 2D Maxwell's equations based on Fourier neural operator," in *Proc. Photon. Electromagn. Res. Symp. (PIERS)*, Nov. 2021, pp. 1635–1643.
[102] B. Zhang and Y. Rahmat-Samii, "On the implementation of Fourier neural operator in solving electromagnetics problems," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Jul. 2022, pp. 693–694.
[103] A. Pietrenko-Dabrowska, S. Koziel, J. W. Bandler, and J. E. Rayas-Sanchez, "EM-driven tolerance optimization of compact microwave components using response feature surrogates," in *IEEE/MTT-S Int. Microwave Symp. Dig.*, Jun. 2022, pp. 107–110.
[104] S. Koziel and A. Pietrenko-Dabrowska, "Tolerance-aware multi-objective optimization of antennas by means of feature-based regression surrogates," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5636–5646, Jul. 2022.
[105] S. Koziel, A. Pietrenko-Dabrowska, and U. Ullah, "Tolerance-aware optimization of microwave circuits by means of principal directions and domain-restricted metamodels," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 9, pp. 4085–4093, Sep. 2022.
[106] J. A. Easum, J. Nagar, P. L. Werner, and D. H. Werner, "Efficient multiobjective antenna optimization with tolerance analysis through the use of surrogate models," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 6706–6715, Dec. 2018.
[107] A. Pietrenko-Dabrowska and S. Koziel, "Low-cost feature-based tolerance optimization of multi-band antennas," in *Proc. 16th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2022, pp. 1–5.
[108] Z. Hamici, "Fast beamforming with fault tolerance in massive phased arrays using intelligent learning control," *IEEE Trans. Antennas Propag.*, vol. 67, no. 7, pp. 4517–4527, Jul. 2019.
[109] W. Chen, Q. Wu, C. Yu, H. Wang, and W. Hong, "Multibranch machine learning-assisted optimization and its application to antenna design," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 4985–4996, Jul. 2022.
[110] J. Zhang et al., "A novel surrogate-based approach to yield estimation and optimization of microwave structures using combined quadratic mappings and matrix transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 8, pp. 3802–3816, Aug. 2022.
[111] L. Y. Xiao, W. Shao, F. L. Jin, B. Z. Wang, and Q. H. Liu, "Inverse artificial neural network for multiobjective antenna design," *IEEE Trans. Antennas Propag.*, vol. 69, no. 10, pp. 6651–6659, Oct. 2021.
[112] T. Q. Duong, L. D. Nguyen, B. Narottama, J. A. Ansere, D. Van Huynh, and H. Shin, "Quantum-inspired real-time optimization for 6G networks: Opportunities, challenges, and the road ahead," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1347–1359, 2022.
[113] C. Wang and A. Rahman, "Quantum-enabled 6G wireless networks: Opportunities and challenges," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 58–69, Feb. 2022.
[114] S. J. Nawaz, S. K. Sharma, S. Wyne, M. N. Patwary, and M. Asaduzzaman, "Quantum machine learning for 6G communication networks: State-of-the-art and vision for the future," *IEEE Access*, vol. 7, pp. 46317–46350, 2019.
[115] J. A. Ansere, T. Q. Duong, S. R. Khosravirad, V. Sharma, A. Masaracchia, and O. A. Dobre, "Quantum deep reinforcement learning for 6G mobile edge computing-based IoT systems," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2023, pp. 406–411.
[116] J. A. Ansere, E. Gyamfi, V. Sharma, H. Shin, O. A. Dobre, and T. Q. Duong, "Quantum deep reinforcement learning for dynamic resource allocation in mobile edge computing-based IoT systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 6, pp. 6221–6233, Jun. 2024, doi: 10.1109/TWC.2023.3330868.