

# Artificial Neural Networks for Microwave Computer-Aided Design: The State of the Art

Feng Feng<sup>1</sup>, Member, IEEE, Weicong Na<sup>2</sup>, Member, IEEE, Jing Jin<sup>3</sup>, Member, IEEE,  
 Jianan Zhang<sup>4</sup>, Member, IEEE, Wei Zhang<sup>5</sup>, Member, IEEE, and Qi-Jun Zhang<sup>6</sup>, Fellow, IEEE  
 (Invited Paper)

**Abstract**—This article presents an overview of artificial neural network (ANN) techniques for a microwave computer-aided design (CAD). ANN-based techniques are becoming useful for performing forward/inverse modeling for active/passive components to enhance a circuit design. With measured or simulated data of microwave devices, ANNs can be trained to learn relevant microwave relationships, which are, otherwise, computationally expensive or for which efficient analytical formulas are not available. Fundamental concepts of the ANN structure and training, such as feedforward neural networks (FFNNs), recurrent neural networks (RNNs)/dynamic neural networks (DNNs)/time-delay neural networks (TDNNs), deep neural networks, and neural network training and extrapolation, are described. Knowledge-based neural networks (KBNNs) are described for improving the accuracy and reliability of modeling and design optimization. Various advanced ANN techniques, such as neuro-transfer function (neuro-TF) modeling, neural network inverse modeling, and deep neural network modeling, are discussed. The existing and emerging applications of ANN in microwave CAD are identified, such as electromagnetic (EM)/multiphysics modeling, modeling of nonlinear circuits and transistors, filter design, very large-scale integration (VLSI) interconnects, oscillator, transmitter and receiver modeling, and CAD applications in such as gallium nitride (GaN) high electron-mobility transistor (HEMT), wireless power transfer (WPT), microelectromechanical system (MEMS), and substrate-integrated waveguide (SIW).

**Index Terms**—Artificial neural networks (ANNs), deep neural network, inverse modeling, knowledge-based neural network (KBNN), microwave computer-aided design (CAD), neuro-transfer function (neuro-TF).

Manuscript received 1 June 2022; accepted 12 July 2022. Date of publication 19 August 2022; date of current version 4 November 2022. This work was supported by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN-2017-06420. (Corresponding author: Feng Feng.)

Feng Feng is with the School of Microelectronics, Tianjin University, Tianjin 300072, China (e-mail: ff@tju.edu.cn).

Weicong Na is with the Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (e-mail: weicongna@bjut.edu.cn).

Jing Jin is with the College of Physical Science and Technology, Central China Normal University, Wuhan 430079, China (e-mail: jingjin@ccnu.edu.cn).

Jianan Zhang is with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China (e-mail: jiananzhang@seu.edu.cn).

Wei Zhang is with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: weizhang13@bupt.edu.cn).

Qi-Jun Zhang is with the Department of Electronics, Carleton University, Ottawa, ON K1S5B6, Canada (e-mail: qjz@doe.carleton.ca).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMTT.2022.3197751>.

Digital Object Identifier 10.1109/TMTT.2022.3197751

## I. INTRODUCTION

ARTIFICIAL neural networks (ANNs) are important techniques for a microwave computer-aided design (CAD) to perform forward/inverse modeling for active/passive components to enhance a circuit design [1], [2], [3], [4], [5]. With measured or simulated data of microwave devices, ANNs can be trained to learn relevant microwave relationships, which are, otherwise, computationally expensive or for which efficient analytical formulas are not available. By training an ANN using data from electromagnetic (EM)/physics simulations, one can use the trained ANN as models for microwave devices to replace the EM/physics models, which are typically CPU-intensive, to significantly accelerate the circuit design with EM/physics-level accuracies.

ANNs can help address two of the frequently encountered challenges in microwave CAD: One is the computationally expensive challenge in forward modeling, and the other is the no-analytical-equation challenge in the inverse design. The types of ANNs used in microwave CAD have ranged from fundamental neural networks, such as multilayer perceptrons (MLPs) and radial basis function (RBF) networks, to advanced neural networks, such as knowledge-based neural networks (KBNNs) for handling reduced amount of training data and deep neural networks for handling a large number of input variables. In this article, we review the state of the art of the ANN techniques for microwave CAD, from fundamental neural networks to advanced KBNN, neural network-based inverse modeling, and deep neural networks.

Early works in ANN for microwave CAD cover a variety of microwave applications, such as microstrip circuit design [6], [7], [8], [9], spiral inductor [10], impedance matching [6], [11], and MESFET devices and circuits [12], [13]. Increased research interests of ANN in the microwave community in the 1990s led to special issues on applications of ANN in microwave CAD (1999 and 2002 special issues of the *Int. J. RF Microwave CAE*) covering more topics, such as ANN structures and training [14], [15], [16], EM design acceleration [17], [18], [19], microwave filter design [17], [20], [21], [22], microwave amplifier design [23], [24], [25], [26], intermodulation distortion [27], coplanar waveguide (CPW) [28], microwave device modeling [29], microstrip antennas [30], [31], large-signal analysis [32], [33], multiconductor transmission line [34], and microwave

measurements [35]. An early review of ANN for microwave CAD is given in [14]. A systematic description of the area from theory to practice was presented in [2]. The work in [3] provided a review of further developments for ANN-oriented EM optimization.

To improve the accuracy and reliability of ANN modeling and design optimization, the KBNN has been developed [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50]. The knowledge-based approach combines neural networks with prior knowledge to build models. Analytical expressions, empirical models, and equivalent circuits are the most commonly used formats of prior knowledge in this kind of approach. It has been shown that using prior knowledge not only accelerates model development but also improves the overall model's learning and generalization ability [51], [52], [53].

The neuro-transfer function (neuro-TF) modeling approach, which integrates neural networks with transfer functions, has emerged as an attractive candidate in EM parametric modeling in recent years [54], [55], [56], [57], [58], [59], [60]. In this method, transfer functions, which have been widely used in EM modeling and optimization [61], [62], [63], [64], [65], [66], [67], are used to represent the highly nonlinear EM responses versus frequency. The remaining less nonlinear relationship between transfer function parameters and design variables, for which no analytical formula is available, is learned by neural networks. Once developed, the neuro-TF model can be subsequently utilized in microwave design optimizations [68], [69], [70], [71], [72].

The ANN has also been trained to learn the complex and high-dimensional relationships between inputs and outputs in the inverse problems [73], [74], [75], [76], [77], [78]. In many situations, the inverse input–output relationship in the training data may have non-uniqueness problems, which means that different training samples with the same input values have contradictory output values. A fundamental issue in this research direction is that, for the same input values, it is impossible to have a trained ANN model match contradictory output values simultaneously. Therefore, it is often challenging to obtain an inverse model with a small training error and high model accuracy. Various advanced inverse modeling techniques [79], [80], [81], [82], [83] have addressed this non-uniqueness problem in the microwave area.

The last decade has witnessed the great research efforts devoted to the development of neural network models with many hidden layers (known as the deep neural network models) in the research community [84], [85], [86], [87], [88], [89], [90]. Deep neural network has been recognized as a very powerful tool in modeling intricate relationships in large datasets [84]. The advantages of deep neural networks are then leveraged to address the challenges in high-dimensional microwave component modeling [91].

In the subsequent part of this article, we first describe the fundamental concepts of ANN structure and training in Section II. The various KBNN methods are discussed in Section III. The formulation and training algorithm of neuro-TF are described in Section IV. The neural network-based inverse modeling algorithms are reviewed in Section V.

In Section VI, the structure and deep learning algorithm of the various deep neural network techniques for microwave modeling are presented. Several major categories of applications of ANN in microwave CAD are highlighted in the review in Section VII. Section VIII concludes this article and discusses the future directions of ANN in microwave CAD.

## II. ANN STRUCTURE AND TRAINING

A neural network consists of a number of neurons and a corresponding set of links connecting between them. Different neuron types or different manners in which the neurons are connected can form different neural network structures [1].

### A. Feedforward Neural Networks

Feedforward neural networks (FFNNs) are ANNs wherein connections between the neurons do not form a cycle. FFNNs are usually used to solve non-dynamic modeling problems [1]. MLPs are the most popularly used FFNN structures, which are widely used in microwave modeling for both passive component modeling [7], [8], [9], [10], [15], [17], [18], [19], [22], [30], [31], [34], [76], [80], [83], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101], [102], [103], [104], [105] and active device/circuit modeling [12], [13], [23], [24], [25], [29], [32], [33], [106], [107], [108], [109], [110], [111], [112], [113], [114], [115], [116], [117], [118], [119], [120], [121]. In the MLP structure, neurons are grouped into different layers. Three types of layers are defined, i.e., input layer (the first layer), output layer (the last layer), and hidden layers (the rest of the layers in between). Another type of FFNN widely used in microwave design is the RBF network [15], [24], [27], [122], [123], [124], [125]. The output of the RBF network is a linear combination of RBFs of the inputs and neuron parameters. The RBF-neural network (RBF-NN) is more effective in representing various localized behaviors in the input–output relationship [126]. When the behavior of the problem exhibits high nonlinear phenomena or contains sharp variations, wavelet neural networks (WNNs) can be used, since the localized nature of their hidden neurons makes it easier to train and obtain a promising model accuracy [15], [21], [24]. As a single-hidden layer FFNN, the extreme learning machine (ELM) is found to have a fast learning speed and good performance in EM parametric modeling when the training dataset is not too large [127], [128], [129].

### B. Dynamic/Recurrent/Time-Delay Neural Networks

A time-domain dynamic behavior of nonlinear devices or circuits is typically characterized by dynamic neural networks (DNNs) [130], [131], recurrent neural networks (RNNs) [26], [132], [133], [134], [135], [136], and time-delay neural networks (TDNNs) [137], [138], [139]. The DNN is a special type of feedback neural network, which uses continuous-time formulation to model the dynamic microwave behavior. The neural network output of DNN is a function of the input–output signals and their time derivatives. RNN is another type of neural network with feedback from output to input. It has the capability of representing a dynamic system's

behavior through a suitable learning process. The RNN uses discrete-time formulation to represent the dynamic behavior of microwave circuits. The neural network output of RNN is a function of its present input and a history of its input and output. The TDNN is an FFNN whose output is formulated by a time-delay series of inputs in discrete-time formulation, such as RNN. TDNN is simpler than RNN, since the output of TDNN is a function of its present input and history of only its input.

### C. Deep Neural Network

A deep neural network is defined as a neural network with many hidden layers [84], [88], [89], [90], [140], [141], [142], [143]. It has been shown that deep neural networks have much better performance than shallow neural networks (i.e., neural networks with a few hidden layers) in dealing with modeling problems exhibiting very intricate relationships. One type of deep neural networks is the deep MLP [89], [91], [142], [144]. The selection of hidden neuron activation functions in deep MLP could affect the ultimate training accuracy. In deep MLP, one of the most popular activation functions is the rectified linear unit (ReLU). The ReLU helps solve one of the main problems, i.e., the vanishing gradient problem, in training deep neural networks. A convolutional neural network (CNN) is another type of deep neural network [77], [145], [146]. A typical CNN is structured as a series of stages. Each stage usually includes three different layers: a convolutional layer, a nonlinearity layer, and a pooling layer [84], [145]. CNN can be used to learn EM problems where the inputs and/or outputs can be treated in image-like representations [74], [146]. A deep belief network (DBN) is also in the class of deep neural networks. In contrast to the FFNNs, such as deep MLP and CNN, DBNs are hybrid networks that contain both directed and undirected layers [147]. DBN has shown advantages in handling the inverse modeling problem for microwave CAD [148]. Another type of deep neural network is the long short-term memory (LSTM) [149], which is an advanced type of RNN. LSTMs can handle the issue of vanishing gradient encountered during the training procedure of traditional RNNs [149]. LSTM has been introduced into the microwave area for nonlinear device modeling [150] and extrapolation of frequency domain EM responses [151].

### D. Knowledge-Based Neural Network

Taking advantage of a wide range of existing equivalent circuit models/empirical models for microwave components, KBNNs have been developed for microwave CAD [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53]. The use of knowledge in neural networks helps to reduce the amount of training data needed and enhance the extrapolation capability of the model. Further description of KBNN will be included in Section III.

An advanced knowledge-based modeling approach, that integrates neural networks with transfer functions (shortened as neuro-transfer function or neuro-TF), can be used to build effective parametric models of EM structures [54], [55], [56], [57], [58], [59], [60]. This knowledge-based approach can

be used even if accurate equivalent circuits or empirical models are unavailable. Further description of neuro-TF will be included in Section IV.

### E. Discussions for Different Neural Network Structures

Different ANN structures can be chosen for different applications. When modeling simple and straightforward relationship and data generation is not expensive, the FFNN is the simplest and most efficient neural network structure to be used. Sufficient amount of training data is the key to ensure the quality of the neural network training and accuracy of the trained neural network. When the data generation is computationally expensive, KBNN becomes more suitable, since the prior knowledge can help decrease the amount of training data needed while maintaining good modeling accuracy. When the equivalent circuit/empirical model is not practically available and the input–output response of the training data behaves a highly nonlinear response with sharp ripples with respect to frequency, neuro-TF can be applied and can be trained with smaller amount of data than pure ANN structures. To address the situation where the dimension of model inputs is high and/or the amount of training data is large, deep neural networks typically become better choices. When the neural network outputs represent time-domain responses, especially having memory effects, DNN/RNN/TDNN structures are needed.

### F. Neural Network Training and Data Generation

Neural networks can only represent a considered microwave device's or circuit's behavior after it is well trained with corresponding data [1]. The neural network training involves the use of training data to guide the adjustment of neural network internal connection weights (or synaptic connection weights). A separate set of data, called the test data, is used to assess the quality of the trained neural network. It is expected that the test data are not used during training, to fairly assess the trained neural network. The process of developing a neural model generally involves six steps. They are as follows.

- Step 1):* Identify inputs and outputs of the microwave modeling problem.
- Step 2):* For the identified inputs and outputs, generate training and testing data.
- Step 3):* Identify possible knowledge of the overall problem or a subproblem. Formulate ANN or knowledge-based ANN structure.
- Step 4):* Perform model training, model validation, and ANN structural adaptation.
- Step 5):* Test the accuracy of the trained model.
- Step 6):* Incorporate the model into microwave CAD.

Data generation is an important issue for ANN model development. Loosely speaking, an ANN model's accuracy enhances, as the number of data samples increases. However, generating large numbers of data samples is inevitably time-consuming and resource-demanding (e.g., 3-D EM simulations). Various data sampling strategies have been introduced for ANN model development, such as uniform/nonuniform

grid distribution [1], star distribution [37], random distribution [152], Latin hypercube sampling (LHS) [153], [154], adaptive sampling [155], [156], [157], and design of experiment (DoE) with orthogonal distribution [55], [70]. In grid distribution, each input parameter is sampled at equal or unequal intervals. In star distribution, sample points are generated by perturbing a central point twice along each input dimension, once toward the positive direction, and once toward the negative direction. When the input dimension increases, the number of data samples using grid distribution grows exponentially, while the number of data samples using star distribution increases linearly. The grid distribution provides the most sufficient data for training, but is not practically usable when the input dimension is high. The star distribution cannot provide sufficient data for pure neural network training, but it can be used for training neural networks combined with prior knowledge and has the least number of training data among all the sampling methods, especially when the input dimension is high. In the random distribution, each input sample is a vector of random variables within the range of input parameters. LHS improves the uniformity of distributed samples in the sampling space compared with random distribution. Adaptive sampling uses the interaction between the intermediate surrogate model and the available data to adaptively and iteratively build an optimized training dataset. It can distinguish where data are needed in the input modeling space. Compared with grid sampling and random sampling, adaptive sampling helps to achieve the required model accuracy using the fewest data samples, making the data generation process flexible and efficient. DoE with orthogonal distribution can be used for generating multiple sample points in high-dimensional parameter spaces. In DoE, the subspace divisions are orthogonal and sampled with the same density. The use of orthogonal distribution around the center point allows DoE to have far fewer sampling points than grid distribution does. Due to this, the surrogate model developed by DoE can be valid in a much larger neighborhood than that developed by star distribution. With the increasing model dimension and complexity, data generation in ANN modeling faces new challenges. How to keep a good balance between the cost of data sampling and the accuracy of a model is an ongoing direction to be further addressed in the future.

A robust automated model generation (AMG) algorithm [158], [159] for neural network training can be used to significantly reduce the intensive human effort demanded by the conventional step-by-step neural modeling approach. The algorithm inherently distinguishes between smooth and nonlinear regions of model behavior. It then uses relatively fewer samples in smooth subregions and more samples in nonlinear subregions. The AMG algorithm can systematically generate a neural network model satisfying the user-desired accuracy. Active learning is widely used in deep learning area and can adaptively perform model generation. In active learning, adaptive sampling is used to create a feedback loop between the deep learning model and the data [160], [161]. As more samples are collected, the adaptive sampling method learns how to improve the sampling process by deciding where to sample next. This enables faster training convergence,

as more data are generated in the regions that have maximum uncertainty. Incorporation of active learning into the AMG process is an interesting further step.

### G. Extrapolation for Neural Networks

Typically, a standard neural network model is accurate only inside the region where it is trained and is unreliable outside this training region. Various extrapolation methods [162] for neural network models have been introduced to address this issue, such as model-level extrapolation [163], [164], neuron-level extrapolation [165], and multidimensional extrapolation [166]. To facilitate the model extrapolation, the model-level extrapolation methods use the neural network and its derivative information at the training region's boundaries [163], [164]. Neuron-level extrapolation method addresses the problem in case when the training boundaries are irregular [165]. This method detects the necessity of extrapolation inside each hidden neuron and performs extrapolation using a modified neuron activation function. The simplification of the problem into a combination of several 1-D first-order extrapolations makes this method simple and easy to be implemented. To address the modeling challenges in EM optimization and large-signal harmonic balance simulation, an advanced multidimensional extrapolation technique for microwave modeling and design has been used [166]. With a good model accuracy being guaranteed inside the training region, the extrapolation technique [166] can further force the model to be smooth along all the directions outside the training region. Besides various extrapolation methods, the KBNN method [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53] exploits microwave empirical or equivalent models to help the overall KBNN model to achieve better extrapolation capability than pure neural networks. Furthermore, CNN- and LSTM-based neural network structures have been exploited for design space extrapolation and frequency extrapolation [151].

## III. KNOWLEDGE-BASED NEURAL NETWORKS

The KBNN incorporates the available microwave information in the form of microwave equivalent circuit/semi-analytical models (as prior knowledge) into neural network structures. There are several ways in which microwave knowledge and neural network are combined, as shown in Fig. 1. Let  $\mathbf{y}$  represent the outputs of the KBNN model. Let  $\mathbf{x}$  represent the inputs of the KBNN model, i.e., geometrical parameters. Let  $\mathbf{d}$  represent the values of  $\mathbf{y}$  in the training or validation data (such as data generated from EM simulation). Fig. 1(a) shows the general structure of KBNN, followed by special structures of various KBNN methods in Fig. 1(b)–(g).

### A. Basic KBNN Structures

The difference method [9] uses ANN to learn the difference between EM data and prior-knowledge model outputs, illustrated in Fig. 1(b). The inputs of ANN are physical parameters, and the outputs are the differences between the prior-knowledge model and the EM simulation. The difference

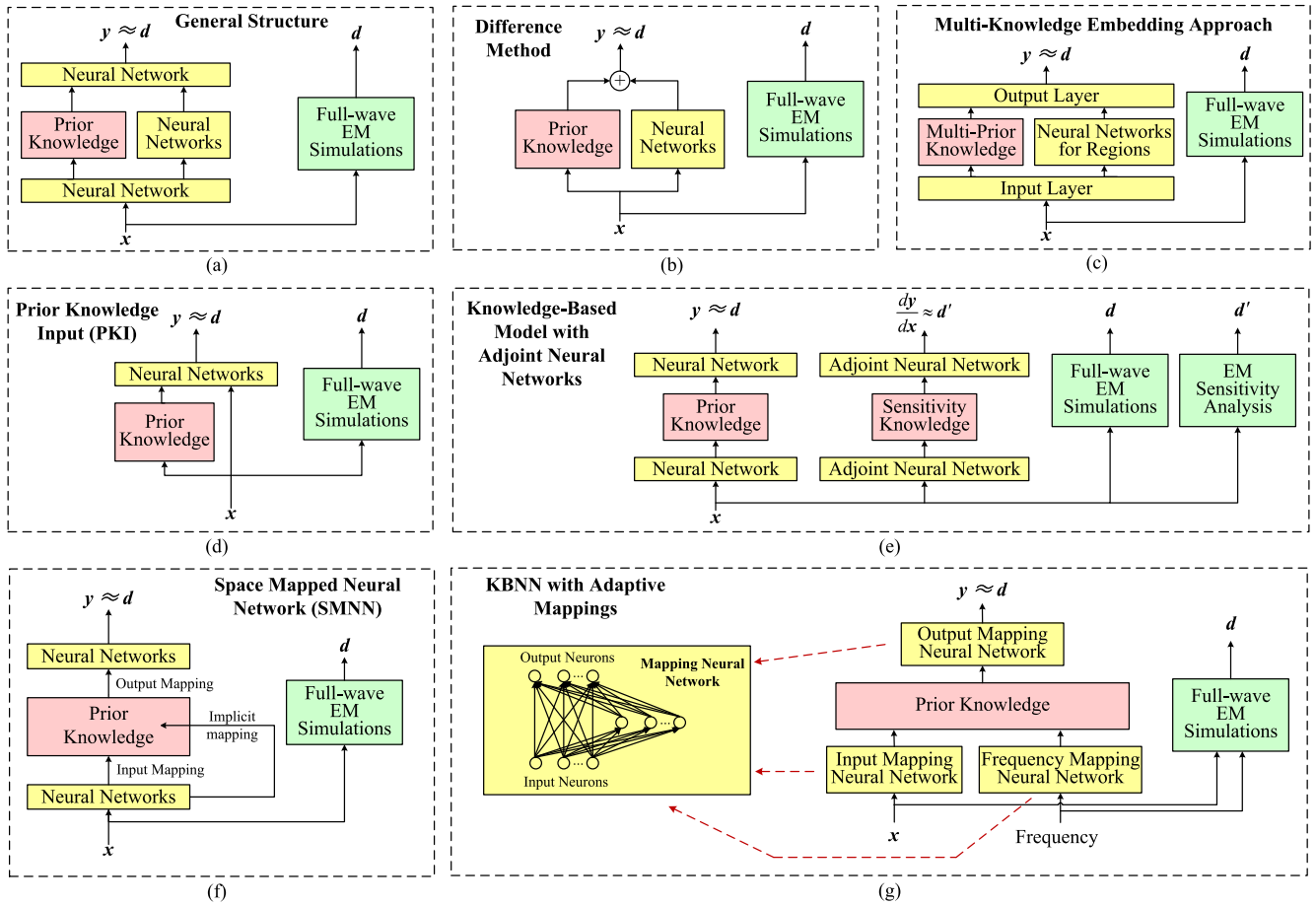


Fig. 1. (a) General structure of KBNN. (b) Structure of the difference method for KBNN [9]. (c) Structure of the multiknowledge embedding approach for KBNN [36]. (d) Structure of the PKI method [28]. (e) Structure of the knowledge-based model with adjoint neural networks [58]. (f) Structure of the SMNN method [37], [167]. (g) Structure of KBNN with adaptive mappings [44].

method is very efficient in dealing with the modeling problem where the differences between the outputs of the prior knowledge and the EM simulation have simpler relationship with respect to the inputs compared with the outputs of the EM simulation itself.

The multiknowledge embedding approach [36] for KBNN embeds the multiple prior knowledge into separate subspaces of the neural network model, resulting in an overall model covering a larger space than any of the individual knowledge models alone, as shown in Fig. 1(c). Compared with the difference method, the multiknowledge embedding approach embeds the prior knowledge into the internal of neural networks, allowing the boundaries between subspaces for each prior knowledge to be learned too.

In the prior-knowledge input (PKI) method [28], the prior-knowledge model outputs are used as the inputs to the ANN, in addition to the original problem (EM data) inputs, as illustrated in Fig. 1(d). The PKI method is more suitable when dealing with the modeling problem where the similarities between prior-knowledge output and EM output lead to simpler training of the ANN.

Space mapped neural networks (SMNNs) [22], [37], [167] concept divides the modeling problem into a coarse model

(prior-knowledge models) and a fine model (EM simulations). ANN is used to map between the fine model and the coarse model, as shown in Fig. 1(f). SMNN concepts can be applied for input mapping, output mapping, and implicit mapping. By exploiting the vast set of available knowledge models, SMNN methods reduce the number of fine model data needed for training, improve the model generalization ability, and simplify the ANN topology with respect to pure ANN modeling methods.

### B. Knowledge-Based Model With Adjoint Neural Networks

To further speed up the KBNN development process, EM sensitivity information is incorporated into the KBNN model [58], as shown in Fig. 1(e). Let  $d'$  represent the values of derivatives of  $y$  with respect to  $x$  in the training or validation data (such as data generated from EM sensitivity analysis). The knowledge-based model with adjoint neural networks [92] allows robust parametric model development by learning both the input–output behavior and the derivatives of the EM modeling problem simultaneously. Compared with the KBNN modeling without EM sensitivity, the same modeling accuracy can be achieved with fewer training data by exploiting EM

sensitivities. The knowledge-based model with adjoint neural networks is only utilized during the training process. The final model for a user is the simple original knowledge-based model, which can be further used in high-level designs.

### C. KBNN With Adaptive Mappings

In KBNN modeling, there are several different ways to map the difference between the knowledge/coarse model and the EM data, such as input mapping, output mapping, frequency mapping, or a combination of any two or three of these mappings [3], and each mapping can be linear or nonlinear. Different modeling problems or different knowledge/coarse models can lead to different types and complexities of the mapping in KBNN. However, a quantitative decision of mapping structure in a KBNN is unknown in advance. To address this issue, KBNN with adaptive mappings [44] helps to systematically determine the type and topology of the mapping structure in KBNN. The structure of the KBNN with adaptive mappings combines the empirical model with one input mapping, one frequency mapping, and one output mapping to include all cases of mappings. To include both linear and nonlinear mapping, each mapping ANN is an expansion of conventional MLP by adding an additional direct connection between the input layer and the output layer, as shown in Fig. 1(g). A two-stage training strategy with  $l_1$  optimization is presented in [44] to train the model. Using the properties of  $l_1$  norms, the training method automatically determines whether to use linear or nonlinear mapping in various parts of the model in the first-stage training and also determines whether a linear mapping, if exists, can be deleted or not in the second-stage training. After the two-stage training, the appropriate mapping structure of the knowledge-based model for the modeling problem is achieved.

### D. Discussions

This section provides a brief qualitative characterization of the advantages and disadvantages of the different kinds of KBNNs reviewed in this article. Table I shows the comparisons of pure neural network and different KBNN methods. A main advantage of KBNN over pure neural network is reduced amount of training data needed. The efficiency of KBNN also depends on the choice of KBNN methods. The difference method uses the difference between the prior knowledge and the EM simulations to form the model, resulting in a smaller range of the ANN outputs and a simpler input–output relationship. This method is the simplest form of KBNN and is expected to give good results when the difference has a simpler input–output relationship as a function of the inputs than the EM simulation data. The multiknowledge embedding approach allows multiple knowledge functions in different regions of the model. Since these knowledge functions are used for some neurons instead of standard activation functions, the KBNN needs to be trained using methods other than the conventional backpropagation. The advantage is its ability to accommodate multiple knowledge models for different subregions of the overall modeling space and allowing the boundaries of subregions to be learned during training. The PKI method

makes direct correction of the output responses of the prior knowledge. Compared with the difference method, the PKI is slightly more complex. However, the PKI is applicable even if the difference between prior knowledge and EM output has a complicated relationship with respect to inputs. Knowledge-based modeling with an adjoint neural network incorporates the EM sensitivity information into the KBNN model development process. Since more information is provided, the knowledge-based modeling with an adjoint neural network needs fewer EM data to achieve an accurate model than the knowledge-based modeling without sensitivity information. Therefore, this method is suitable if the data generation is very expensive, and sensitivity information is available from the data generator. SMNN uses input mapping, output mapping, or implicit mapping for different microwave modeling problems to construct mapping functions from fine model (EM simulations) to coarse model (the prior knowledge), modify the coarse model, or align the responses between the fine model and the coarse model. When the mapping structure in the SMNN for a microwave modeling problem cannot be determined in advance, KBNN with an adaptive mapping method can be applied to automatically determine the optimal mapping structure for the specific modeling problem. Instead of sequential trials of linear and nonlinear mappings every time for every mapping structure, KBNN with an adaptive mapping method effectively increases the efficiency of knowledge-based model development.

### E. Application Example of KBNN With Adaptive Mappings for Microwave Modeling

As an example, the method of KBNN with adaptive mappings is applied to develop a parametric model of a microwave low-pass filter with double microstrip sections [44], as shown in Fig. 2(a). The geometrical input parameters of the model are the lengths and widths of the microstrip lines of the filter. The output of the model is the magnitude of  $S_{21}$ . Fig. 2(b) shows the prior-knowledge model, which is an equivalent circuit of the low-pass filter using simple transmission lines. For comparison purposes, KBNN models are developed for two different cases. In Case 1, the range of the geometrical input parameters is small, while in Case 2, the range is large. Using the modeling method of KBNN with adaptive mappings, a linear input mapping is good enough to achieve a KBNN with 1.92% testing error for Case 1. For Case 2, the modeling method chooses a nonlinear input mapping together with a nonlinear frequency mapping to achieve a KBNN with 1.97% testing error [44]. From these two modeling examples, it is demonstrated that the modeling method of KBNN with adaptive mappings can adaptively determine the mapping structure of the KBNN model for specific modeling problems. Fig. 2(c) shows modeling results at four different geometrical values for the low-pass filter example.

## IV. NEURO-TF MODELING METHODS

As an advanced knowledge-based modeling approach, the neuro-TF method introduces transfer functions to perform the parametric modeling of the frequency response of microwave

TABLE I  
COMPARISONS BETWEEN PURE NEURAL NETWORK AND DIFFERENT KINDS OF KBNNs

Model	Requirement of the Prior Knowledge	Position of the Prior Knowledge	No. of Neural Networks	Complexity of the Model Structure	EM Sensitivity Information	Mapping Structure Adjustment	Amount of Training Samples
Pure Neural Network	No	-	One	Simple	Not Included	-	Large
Difference Method [9]	Yes	Output	One	Simple	Not Included	No	Small or Medium
Multi-Knowledge Embedding Approach [36]	Yes	Inside	One	Medium	Not Included	No	Small or Medium
Prior-Knowledge Input(PKI) Method [28]	Yes	Input	One	Simple	Not Included	No	Small or Medium
Knowledge-Based Model with Adjoint Neural Network [58]	Yes	Input	One	Complex	Included	No	Small or Medium
Space Mapped Neural Network (SMNN) [167]	Yes	Input or Output or Inside	Multiple	Medium	Not Included	No	Small or Medium
KBNN with Adaptive Mappings [44]	Yes	Input or Output	Multiple	Complex	Not Included	Yes	Small or Medium

components [54], [55], [56], [57], [58], [59], [60], [68], [69], [70], [71], [72]. To develop the neuro-TF model, a transfer function is needed as the prior knowledge to represent the highly nonlinear relationship between EM responses and frequency. Using the knowledge of transfer functions, the remaining relationships of the transfer function parameters and geometrical parameters are less nonlinear and much easier for the neural networks to learn, resulting in the training of neuro-TF to be much more efficient.

Let the order of transfer functions be denoted as  $N$ . Let  $\mathbf{p}$  be a vector containing all the transfer function parameters (e.g., rational function coefficients, or pole/residues, or gain/pole/zeros), i.e.,  $\mathbf{p} = [p_i]_{i=1}^{2N} = [p_1 \ p_2 \ \cdots \ p_{2N}]^T$ . Neural network functions are used to represent the relationships of the transfer function parameters  $\mathbf{p}$  versus  $\mathbf{x}$  and  $\mathbf{w}$ , where  $\mathbf{w}$  represents a vector containing the weights in neural network. The neuro-TF model output in general format is formulated as

$$y(\mathbf{x}, \mathbf{w}, s) = H(\mathbf{p}(\mathbf{x}, \mathbf{w}), s) \quad (1)$$

where  $H$  is the transfer function response;  $s$  is the Laplace domain frequency. Fig. 3 shows the structure of the neuro-TF model in general format.

#### A. Neuro-TF in Different Transfer Function Formats

The neuro-TF in rational format consists of neural networks and rational transfer functions. Let the neural network outputs

represent the coefficients of the rational transfer function. The rational transfer function for neuro-TF is relatively simple. However, its order of the transfer function or its frequency bandwidth is limited to preserve neuro-TF accuracy [54], [59].

Another structure is the neuro-TF model in a pole/residue format, which uses the pole-residue-based transfer function [55] instead of the rational transfer function. The neuro-TF model in the pole/residue format consists of neural networks and pole/residue-based transfer functions. The neuro-TF in pole/residue format has low sensitivities of transfer function parameters with respect to geometrical parameters even when the order of transfer function is high, which can result in a good modeling accuracy and robustness.

To extract the feature parameters from the neuro-TF model to assist EM design optimization, the neuro-TF in pole/zero format is introduced [56], [68], [69]. The neuro-TF in pole/zero format can efficiently extract the transfer function zeros through a calculation to obtain the feature parameters (e.g., feature frequencies) [68], which assists neuro-TF-based optimizations to have better chance of jumping out of local minima [69].

Table II summarizes the formulations and comparisons for the neuro-TF model in different transfer function formats. When the transfer function with relatively low order can represent the EM frequency responses of microwave applications, the neuro-TF in rational format is efficient to use, because the rational coefficients are automatically sorted correctly among

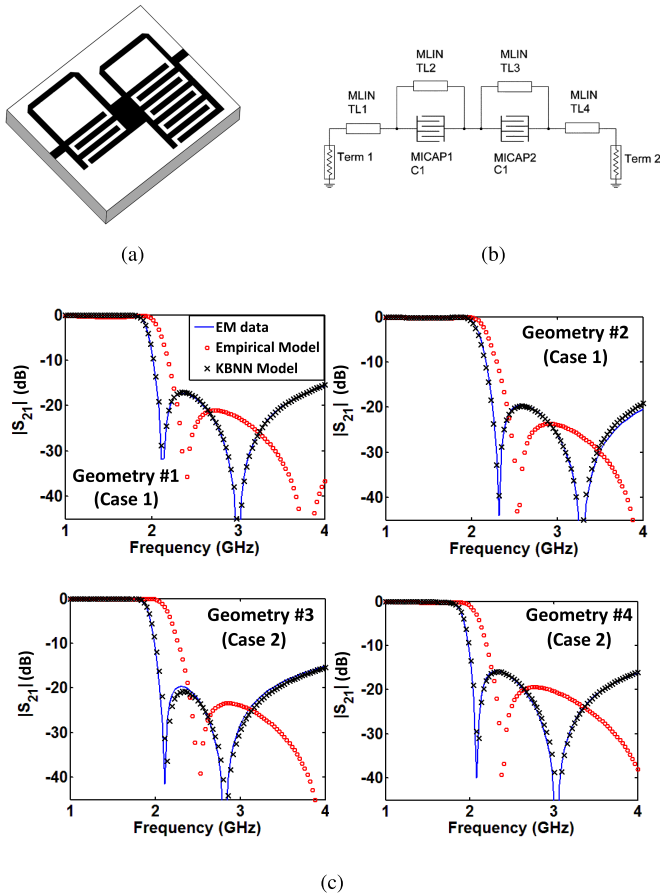


Fig. 2. (a) EM structure of the low-pass filter example. (b) Equivalent circuit model of the low-pass filter example. (c) KBNN modeling results of the low-pass filter example at four different geometrical values for Cases 1 and 2.

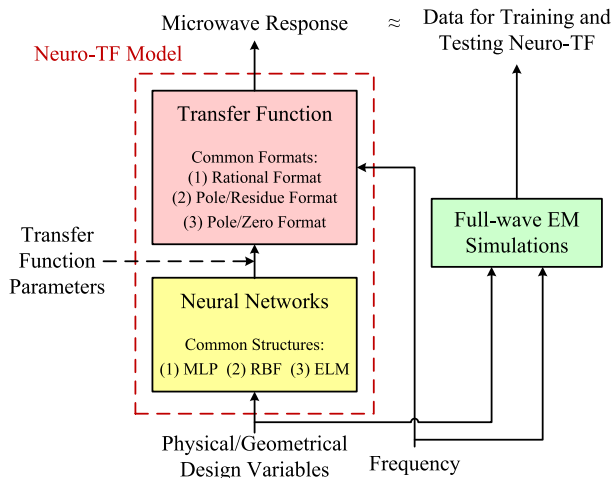


Fig. 3. Structure of the neuro-TF model in general format [55].

different geometrical samples. The neuro-TF model in rational format can automatically avoid the mismatch issue (i.e., different sequences of transfer function parameters among different geometrical samples) in pole/residue or pole/zero format. When the transfer function with relatively high order is needed to represent the EM frequency responses of microwave applications, the neuro-TF model in pole/residue format can achieve

better accuracy and robustness, because the relationship of the response of transfer functions versus the pole/residues is much less sensitive than that versus rational coefficients. When feature parameters are needed to be extracted from a neuro-TF model to assist the design optimization, the neuro-TF model in pole/zero format is effective and efficient to be used.

As an alternative way to model the frequency response of microwave components apart from neuro-TF, a discrete frequency ANN approach can be used to consider each frequency point as a separate output dimension [151], [168]. By removing the frequency variable and converting the information to discretized output format, the training of the neural network becomes simplified. However, it may result in the loss of continuity in the frequency variable, and the dimension in the ANN outputs becomes high when the frequency range increases. Further steps, such as frequency interpolation and frequency extrapolation, would be applied when necessary.

### B. Neuro-TF With Different Neural Network Structures

In the abovementioned neuro-TF methods, different formats of transfer function are used to develop accurate parametric models for different applications, while MLPs are used as the neural network structures [54], [55], [56].

Besides MLPs, different neural network structures are also introduced in neuro-TF methods to expand the methods into broader applications. One of the alternative neural network structures suitable for neuro-TF method is the ELM [127], which fixes input weights and hidden layer bias of a single-hidden layer FFNN. Furthermore, a dynamic adjustment kernel ELM is also used as the neural network structure, which contains increased learning, reduced learning, and hybrid learning in the interior structure [128]. The neuro-TF model with ELM can speed up the training process over that with MLP.

Another alternative neural network structure suitable for the neuro-TF method is the RBF-NN. For specific training of the RBF-NN in neuro-TF, a semi-supervised training, including both supervised and unsupervised training, is implemented [123]. By incorporating the semi-supervised training, the modeling accuracy and efficiency can be improved with the same number of training data.

### C. Two-Stage Training Process for Neuro-TF Models

In general, the neuro-TF model is trained using a two-stage training process. Before training the neuro-TF model, vector fitting [169] is used to extract the transfer function parameters (rational coefficients, or pole/residues, or gain/pole/zeros) from EM responses for all the training samples. The training of the neuro-TF uses a two-stage training process [55]. The first training stage is the preliminary training process, where the relationships of transfer function parameters versus geometrical parameters are learned by neural networks. The second training stage is the model refinement of the overall neuro-TF model. The model refinement is performed by optimizing the weighting parameters in neural networks to minimize the



TABLE II  
COMPARISONS FOR NEURO-TF MODELING METHODS IN DIFFERENT TRANSFER FUNCTION FORMATS

Transfer Function Formats	Neuro-TF Formulations	Complexity of Handling Potential Mismatch in Transfer Function Parameters	Transfer Function Parameter Sensitivities	Convenience in Extraction of Feature Parameters
General Format	$y(\mathbf{x}, \mathbf{w}, s) = H(\mathbf{p}(\mathbf{x}, \mathbf{w}), s)$	-	-	-
Rational Format [54]	$y(\mathbf{x}, \mathbf{w}, s) = \frac{\sum_{i=1}^N p_{N+i}(\mathbf{x}, \mathbf{w}) s^{i-1}}{1 + \sum_{i=1}^N p_i(\mathbf{x}, \mathbf{w}) s^i}$	Simple	High	Inconvenient
Pole/Residue Format [55]	$y(\mathbf{x}, \mathbf{w}, s) = \sum_{i=1}^N \frac{p_{N+i}(\mathbf{x}, \mathbf{w})}{s - p_i(\mathbf{x}, \mathbf{w})}$	Complex	Low	Inconvenient
Pole/Zero Format [69]	$y(\mathbf{x}, \mathbf{w}, s) = p_{2N}(\mathbf{x}, \mathbf{w}) \frac{\prod_{i=1}^{N-1} s - p_{N+i}(\mathbf{x}, \mathbf{w})}{\prod_{i=1}^N s - p_i(\mathbf{x}, \mathbf{w})}$	Complex	Medium	Convenient

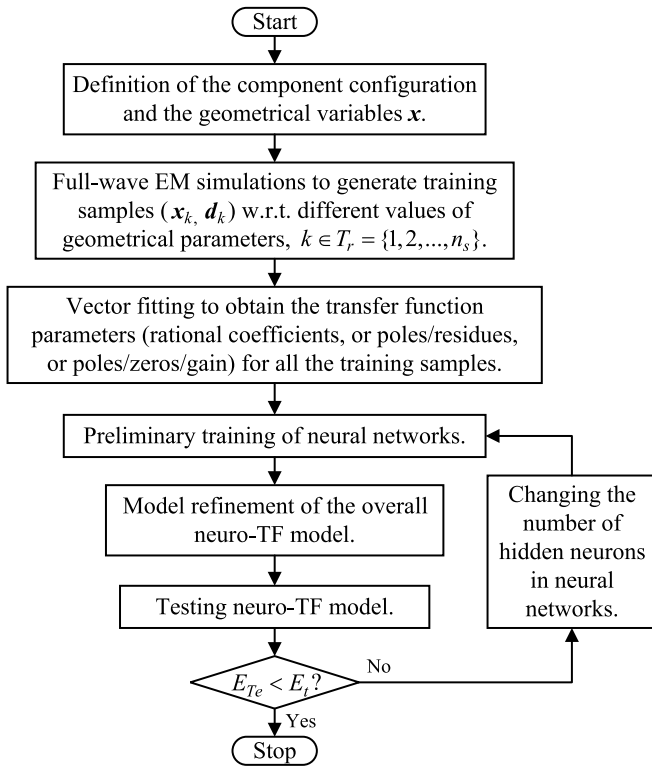


Fig. 4. Flowchart of the overall neuro-TF model development process [55].

training error, denoted as  $E$ , which is formulated as

$$E(\mathbf{w}) = \frac{1}{2n_s n_f} \sum_{k=1}^{n_s} \sum_{j=1}^{n_f} |y(\mathbf{x}_k, \mathbf{w}, s_j) - d_{k,j}|^2 \quad (2)$$

where  $n_s$  represents the number of training samples of various geometrical parameters;  $n_f$  represents the number of frequency samples;  $d_{k,j}$  is the training data generated using EM simulators at the  $k$ th sample and the  $j$ th frequency. Fig. 4 illustrates the flowchart of the development process for the neuro-TF model [55].

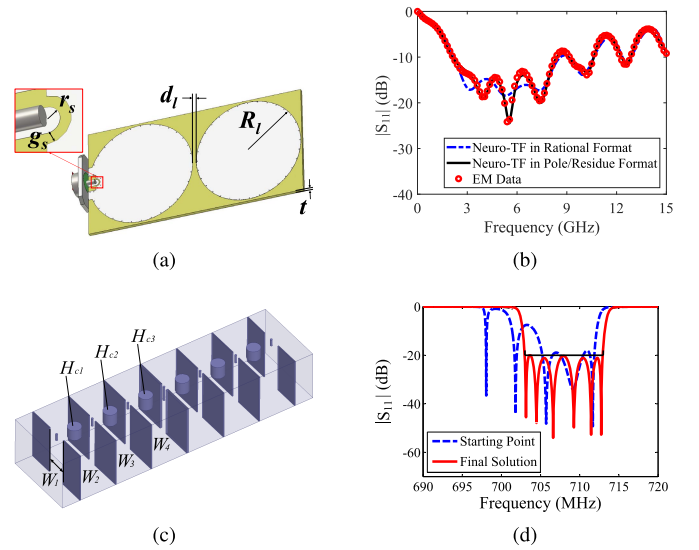


Fig. 5. Application examples for the neuro-TF model. (a) Structure of the UWB antenna [55]. (b) Comparison of the magnitude in decibels of  $|S_{11}|$  of the models developed using different neuro-TF methods and EM data for the UWB antenna [55]. (c) Structure of the sixth-order cavity filter [69]. (d) EM responses of the starting point and final solution of the design optimization using feature-assisted neuro-TF in pole/zero format for the sixth-order cavity filter [69].

#### D. Application Examples for Different Neuro-TF Models

The first example illustrates the development of the neuro-TF model for parametric modeling of the EM behavior of an ultrawideband (UWB) antenna [55], shown in Fig. 5(a), where  $R_l$  and  $d_l$  are the radius of and the distance between the two large pads, respectively.  $t$  is the substrate thickness.  $g_s$  is the gap between the small and large pads connecting to the port.  $r_s$  is the small pad radius. According to the vector fitting results, the order of the transfer function is high in this example. Therefore, the neuro-TF in pole/residue format is suggested. Fig. 5(b) shows that neuro-TF in pole/residue format can achieve better accuracy than that in rational format, since the sensitivities of transfer function output with respect to pole/residues are higher than that with respect to rational coefficients.

For the second example, we perform parametric modeling and design optimization using neuro-TF of the sixth-order cavity filter [69], whose structure is presented in Fig. 5(c). The neuro-TF model has seven geometrical variables as its inputs, i.e.,  $\mathbf{x} = [H_{c1} H_{c2} H_{c3} W_1 W_2 W_3 W_4]^T$ , and one output, i.e.,  $y = |S_{11}|$ . Since feature parameters are needed to be extracted to assist the design optimization in this example, neuro-TF in pole/zero format is suggested. As shown in Fig. 5(d), the design optimization using feature-assisted neuro-TF in pole/zero format successfully achieves the final EM optimal solution, which satisfies the design specifications.

## V. NEURAL NETWORK-BASED INVERSE MODELING

We start the description of inverse modeling by contrasting it to forward modeling. A neural network-based forward model is a neural network trained to represent the original microwave device where the inputs are the physical or geometrical parameters, and the outputs are the electrical parameters. Let  $n$  and  $m$  represent the number of inputs and outputs of the forward model, respectively. The forward model is expressed as

$$\mathbf{y} = \mathbf{f}_{\text{ANN}}(\mathbf{x}, \mathbf{w}) \quad (3)$$

where  $\mathbf{x}$  represents an  $n$ -vector containing the inputs of the forward neural network model;  $\mathbf{y}$  represents an  $m$ -vector containing the outputs;  $\mathbf{f}_{\text{ANN}}$  represents ANN functions for the forward model;  $\mathbf{w}$  represents a vector containing neural network weights.

For the parameter extraction or design purposes of microwave devices, the information is often processed in the reverse way to find the physical or geometrical parameters from desired outputs, which is called the inverse problem. The inverse problem usually can be solved either by an inverse modeling method or an optimization method. For inverse modeling, the model inputs are the electrical parameters, and the model outputs are the physical or geometrical parameters. After a good inverse model is developed, it can obtain the physical or geometrical parameters immediately. This makes the inverse modeling much faster than the iterative optimization process. ANNs have been employed as useful tools for inverse modeling of microwave components [73], [74], [75], [76], [77], [78], [148], [170]. In recent years, various advanced neural network-based inverse modeling techniques have been developed for microwave CAD to address the challenging non-uniqueness problem for inverse modeling [79], [80], [81], [82], [83], [171].

### A. Formulation of Neural Network-Based Inverse Model

Since the inverse input–output relationship is usually unknown and hard to be analytically formulated, a neural network becomes a logical choice to develop the inverse model by learning the input–output relationship based on the training data. For the inverse model, let  $\bar{\mathbf{x}}$  represent an  $m$ -vector containing the inputs, where  $m$  is the number of inputs. The inputs are usually the electrical parameters. Let  $\bar{\mathbf{y}}$  represent an  $n$ -vector containing the outputs, where  $n$  is the number of outputs. The outputs are usually the physical

or geometrical parameters. The neural network-based inverse model is expressed as

$$\bar{\mathbf{y}} = \mathbf{f}_{\text{ANN}}^{-1}(\bar{\mathbf{x}}, \mathbf{w}) \quad (4)$$

where  $\mathbf{f}_{\text{ANN}}^{-1}$  represents the ANN functions for the inverse modeling.

### B. Non-Uniqueness Problem in Inverse Modeling

The training data are generated from the forward model (or simulation of forward problem). The input–output data are then swapped, such that the input data from the forward model are used as the output data during the inverse model training process. When the input–output relationship is unique, i.e., the inverse problem is simple, the training of such an inverse model is easy, and the model is accurate.

For general situations, there will be the non-uniqueness problem of the inverse input–output relationship in the training data. The non-uniqueness problem happens when the different training samples with the same input values have multivalued solutions (contradictory output values). The ANN model cannot learn the contradictory output values simultaneously for the same input values, which results in large training error and low model accuracy. This is the reason why training an inverse model is typically much more challenging than training a forward model.

### C. Derivative-Based Training Method of Neural Network for Inverse Modeling

The derivative-based training method of neural network for inverse modeling addresses the problem of non-uniqueness by dividing the data containing contradictory output values into groups, such that the data in each divided group do not have the problem of contradictory output values [80]. The training data are divided into groups according to the derivatives of outputs versus inputs of the forward model. A neural network inverse submodel is trained for each group of training samples. Since there is no contradictory output values for each group, the training of the subinverse model is simple, and the training accuracy can be high.

After submodels training, the multiple inverse submodels need to be combined to represent the overall model. For this purpose, an algorithm is developed to select the right one among multiple inverse submodels for a given input  $\bar{\mathbf{x}}$ . This is achieved with the help of the forward model. Fig. 6 depicts the inverse submodel combining method for a two submodel system [80]. The input  $\bar{\mathbf{x}}$  is given to each inverse submodel, and outputs from them are fed into accurately trained forward models, respectively, which generate different values of  $\mathbf{y}$ . These outputs of forward models are compared with the input data  $\bar{\mathbf{x}}$ . For the given input, the correct inverse submodel is determined by the least error between  $\mathbf{y}$  and  $\bar{\mathbf{x}}$ .

### D. Multivalued Neural Network Inverse Modeling Technique

In [83], a simpler and more automated multivalued neural network inverse modeling technique is developed to solve the non-uniqueness problem and to simplify the inverse modeling process. This technique addresses the inverse problem

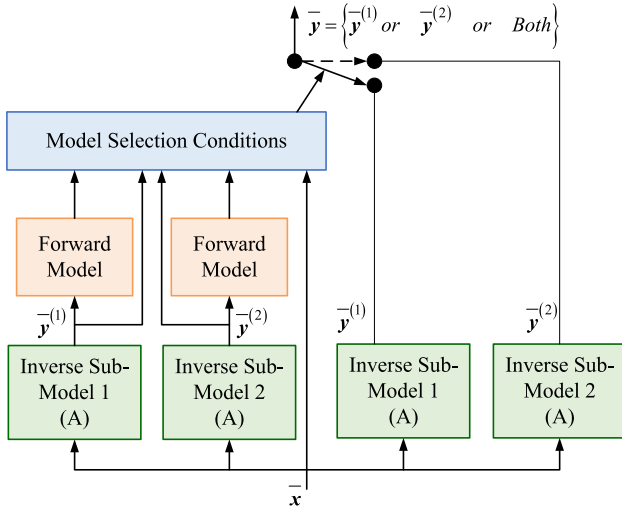


Fig. 6. Neural network inverse model from the derivative-based training method [80].

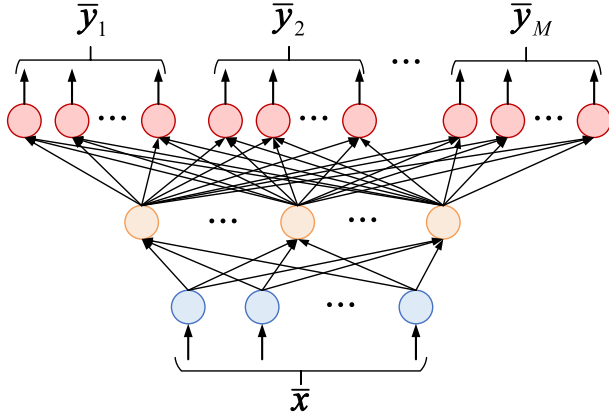


Fig. 7. Multivalued neural network inverse model [83].

of a single set of electrical parameters with multiple sets of physical or geometrical parameters. One set of physical or geometrical parameters is defined as one value of the multivalued inverse model. Let  $M$  represent the number of values of the multivalued inverse model. The overall inverse model is developed by repeating the outputs of the direct inverse model  $M$  times, formulated as

$$[\bar{y}_1^T \quad \bar{y}_2^T \quad \cdots \quad \bar{y}_M^T]^T = \mathbf{g}_{\text{ANN}}(\bar{\mathbf{x}}, \mathbf{w}) \quad (5)$$

where  $\bar{y}_i$  is the  $i$ th value of the multivalued inverse model.  $\mathbf{g}_{\text{ANN}}$  represents the ANN functions of the multivalued inverse model. Fig. 7 shows an example of the multivalued inverse model.

Let  $T_r$  be the index set of all the training samples. Let  $(\bar{\mathbf{x}}_k, \mathbf{d}_k)$ , where  $k \in T_r$ , represent the  $k$ th training sample pair. To solve the non-uniqueness problem, such that different values  $\bar{y}_i$  can match different conflicts for the same or very similar input  $\bar{\mathbf{x}}$ , a training error of the multivalued neural

network is used as [83]

$$E(\mathbf{w}) = \sum_{k \in T_r} E_k(\mathbf{w}) \quad (6)$$

where  $E_k(\mathbf{w})$  represents the training error for the  $k$ th sample, which is defined as

$$E_k(\mathbf{w}) = \left( \sum_{i=1}^M \frac{1}{e_{i,k}(\mathbf{w})} \right)^{-1} \quad (7)$$

where  $e_{i,k}(\mathbf{w})$  represents the error between  $\bar{y}_i$  and  $\mathbf{d}_k$  as follows:

$$e_{i,k}(\mathbf{w}) = \frac{1}{2} \|\bar{y}_i - \mathbf{d}_k\|^2 \quad (8)$$

where  $\bar{y}_i$  is the  $i$ th value of the inverse neural network model  $\mathbf{g}_{\text{ANN}}(\bar{\mathbf{x}}_k, \mathbf{w})$  for the input  $\bar{\mathbf{x}}_k$ . The error function focuses on matching the training sample with one of the values of the multivalued inverse model, which is closest to the current training sample, while ignoring other values. In this way, the non-uniqueness problem is solved, and the developed multivalued inverse model can systematically fit multiple contradictory samples.

### E. Invertible Neural Nets for Inverse Modeling

Invertible neural network models are also developed to handle the inverse modeling problem [171]. Invertible neural networks consist of reversible blocks. In the forward process, the reversible block can compute the output values from input values. In the inverse process, the reversible block can also compute the input values from output values without bringing additional computing complexity compared with the forward process. The model structure of bijectivity of the invertible neural network is suitable for bidirectional operation and training. Therefore, both forward and inverse processes can be well addressed.

The invertible neural network is trained in a bidirectional way in each training iteration. The invertible neural network learns the transformation between the sample distribution and the prior distribution of a latent variable. After the training process, the conditional posterior distribution of the design parameters can be predicted for the given target. This makes this technique especially advantageous for solving the non-uniqueness and degeneracies when multiple candidate solutions exist in the inverse problem.

### F. Discussions

Table III compares different neural network-based inverse modeling methods. Here, we discuss the advantages and situations of using different inverse modeling methods. The simplest ANN method for inverse modeling is called a direct inverse modeling method, where an ANN model is trained using data obtained by swapping the input and output data from the forward problem. When the input–output relationship is unique, i.e., the inverse problem is simple, the training of such an inverse model is easy, and the model is accurate. When the inverse problem is complex and challenging due to the nonuniqueness problem, the derivative-based training

method, the multivalued neural network method, or the invertible neural network method is needed. The derivative-based method addresses the nonuniqueness problem by dividing the inverse model into different inverse submodels according to derivative criteria. Therefore, the derivative-based method requires data preprocessing. If the data preprocessing for the modeling problem is easy, the derivative-based training method is efficient to be used. If the data preprocessing is very intensive and complex, especially in multidimensional input space, the multivalued neural network is more efficient to be used, because it deals with the nonuniqueness problem without the need of data preprocessing. To address both forward and inverse modeling problems simultaneously, the invertible neural network model can be used due to its bidirectional ability.

## VI. DEEP NEURAL NETWORK TECHNIQUES FOR MICROWAVE MODELING

Neural networks with one or two hidden layers have been used to successfully solve many application problems in microwave CAD. Further advance in microwave CAD led to the use of deep neural networks with many hidden layers to address more complex microwave modeling problems. Various deep neural network methods have been explored for microwave CAD, such as deep MLP [89], [90], [91], [142], RNN [133], [140], [150], and CNN [74], [77], [85], [151], [172].

### A. Hybrid Deep MLP for Microwave Modeling

A fundamental type of deep neural network is deep MLP. The number of hidden layers in a deep MLP should be more than three (including three). Training a deep MLP is hard due to the vanishing gradient problem [141]. A hybrid deep MLP can be used for microwave modeling to further solve the vanishing gradient problem and maintain the smooth outputs and continuous derivatives simultaneously [91]. The hybrid deep MLP technique [91] for microwave modeling uses both sigmoid functions and the smooth ReLUs as activation functions, so that the vanishing gradient problem [141] can be overcome, and the number of hidden neurons can be reduced. The smooth ReLU is formulated as [91]

$$f_s(\gamma) = \begin{cases} \gamma, & \gamma > \alpha \\ \frac{1}{4\alpha}\gamma^2 + \frac{1}{2}\gamma + \frac{1}{4}\alpha, & -\alpha \leq \gamma \leq \alpha \\ 0, & \gamma < -\alpha \end{cases} \quad (9)$$

where  $\alpha$  is a value greater than zero.

The structure of the hybrid deep MLP model is shown in Fig. 8 [91]. Let  $q_s$  represent the number of hidden layers using sigmoid functions, which are the layers close to the input layer, as shown in Fig. 8. Let  $q_r$  represent the number of hidden layers using smooth ReLUs, which are the layers close to the output layer, as shown in Fig. 8. The total number of layers in the hybrid deep MLP is expressed as  $q_s + q_r + 2$ .

To deal with the vanishing gradient problem that may exist during training of the hybrid deep MLP, a three-stage deep learning algorithm can be used [91]. In Stage I, the deep

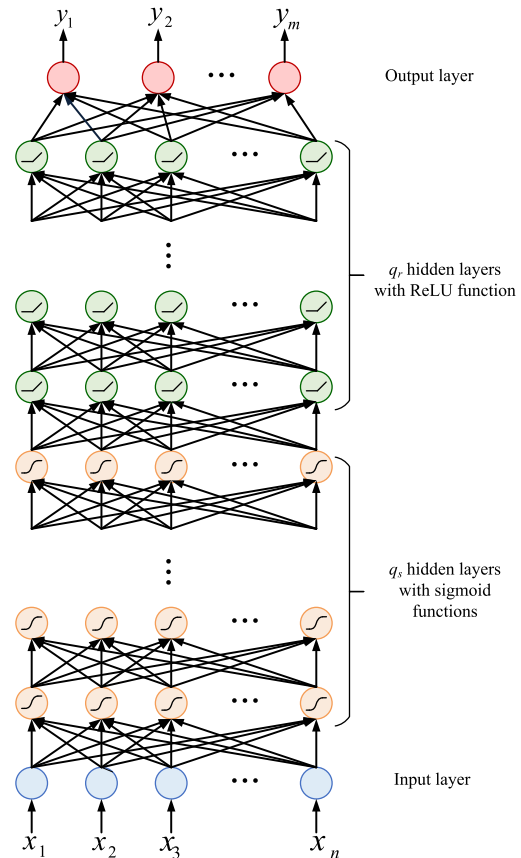


Fig. 8. Structure of the hybrid deep MLP [91].

MLP with  $q_s$  sigmoid hidden layers is trained. Usually,  $q_s$  is initialized as  $q_s = 3$ . New sigmoid hidden layers are added until the vanishing gradient problem begins to appear or until the model accuracy satisfies the requirement. If the model accuracy satisfies the requirement, the training is finished. Otherwise, proceed to Stage II. In Stage II, hidden layers with conventional non-smooth ReLUs are added until the accuracy of the deep MLP model satisfies the requirement. In Stage III, all the conventional non-smooth ReLUs in the trained deep MLP model are replaced by the smooth ReLUs. Then, the hybrid deep MLP with sigmoid functions and smooth ReLUs is further trained to refine the model accuracy.

### B. Local-Feedback Deep RNN

To model the dynamic behavior of a nonlinear microwave device, the RNNs are explored. The RNN structure contains feedback loops involving delay units. Due to the feedback loops, RNN structures are naturally deep in time dimension. This is because the current state of RNN is related to all historic states. An example of a deep RNN is the local-feedback deep RNN (LFDRNN) developed to capture more complex time dependencies [133]. The LFDRNN structure consists of multiple local-feedback RNN layers. In the LFDRNN structure, time-delayed local feedbacks are incorporated to improve the convergence of the RNN model [173].

TABLE III  
COMPARISONS OF DIFFERENT INVERSE MODELING METHODS

Inverse Modeling Methods	Addressing Nonuniqueness Problem	No. of Neural Networks	Complexity of Neural Network Training	Complexity of Data Preprocessing	Human-Intensiveness When Addressing Nonuniqueness Problem	Bi-Directional Modeling Ability
Direct Inverse Modeling method [18]	No	One	Simple	Simple	-	No
Derivative Based Training Method [80]	Yes	Multiple	Simple	Complex	Much	No
Multivalued Neural Network [83]	Yes	One	Complex	Simple	Little	No
Invertible Neural Network [171]	Yes	One	Complex	Simple	Medium	Yes

Similar to other ANN structures, the training process of the LFDRNN is a process to optimize the objective function by adjusting the neural network weight parameters. To train the LFDRNN, the gradients of model outputs with respect to weights should be computed both through time and through layer weights [133]. By introducing the time-delayed local feedbacks, the LFDRNN can improve the ability of modeling the complex nonlinear device behaviors as compared with the conventional RNNs.

### C. Augmented CNN

An example of CNN for microwave CAD is the augmented CNN, developed for behavioral modeling and digital predistortion of concurrent multiband power amplifiers [172]. To reduce the model complexity and improve the linearization performance of the digital predistortion model, the augmented CNN is composed of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer. In the input layer, the input signals are reshaped into 2-D matrices. The layer following the input layer is the convolutional layer. For the purposes of convolution, sliding convolutional filters are employed. Most of the computations of the CNN structure are performed by the convolutional layer. The convolved outcomes are added with the biases. The convolutional layer is followed by the pooling layer, which is a many-to-one mapping used to decrease the size of the output matrix from the convolutional layer. The pooling layer can also reduce the risk of overfitting during training. The pooling layer is followed by the fully connected layer. The fully connected layer is followed by the output layer, where linear functions are used as the activation functions.

### D. Discussions

Table IV compares the different deep neural network techniques. Here, we discuss the advantages and situations of using different deep neural network techniques. If the dimension of model inputs is relatively low, the shallow MLP (i.e., MLP with one or two hidden layers) is easy to use for modeling. When the dimension of model inputs becomes high, the deep MLP is needed for accurate modeling. When the

training of a deep MLP becomes hard due to the vanishing gradient problem, a hybrid deep MLP can be used to solve the vanishing gradient problem and maintain the smooth outputs and continuous derivatives simultaneously. Since the hybrid deep MLP is a fully connected neural network, the connections between neurons are easy to establish, and the training complexity is relatively low. Compared with the deep MLP, the CNN is a partially connected neural network where the connections are established between subgroups of neurons. The size of the subgroups and the connection patterns between the subgroups affect the structure of the CNN model. Therefore, the CNN has a better ability to extract features/patterns than deep MLP. Besides, CNN is specifically good at addressing modeling problems with image-like inputs. For the time-domain modeling with memory effects, the RNN is effective to use, since the model outputs are the functions of the present and past time series of model inputs and outputs. Compared with the conventional RNN with depth only in the time dimension, the LFDRNN can capture more complex time dependencies, because it has depth both in the time and in the space dimensions.

### E. Application Example of the Hybrid Deep MLP for Microwave Modeling

In this example, the hybrid deep MLP technique is used to develop a parameter-extraction model for the sixth-order multicoupled cavity filter. The center frequency of the filter is 11785.5 MHz, and the bandwidth is 56.2 MHz [174]. The ideal coupling matrix for this filter is

$$M_{\text{ideal}} = \begin{bmatrix} -.0473 & .8489 & 0 & 0 & 0 & 0 \\ .8489 & -.0204 & .6064 & 0 & 0 & 0 \\ 0 & .6064 & -.0305 & .5106 & 0 & -.2783 \\ 0 & 0 & .5106 & .0005 & .7709 & 0 \\ 0 & 0 & 0 & .7709 & -.0026 & .7898 \\ 0 & 0 & -.2783 & 0 & .7898 & .0177 \end{bmatrix}. \quad (10)$$

The outputs of the deep MLP model are the 12 nonzero coupling parameters in the  $M_{\text{ideal}}$  matrix, i.e.,  $\bar{y} = [M_{11} M_{22} M_{33} M_{44} M_{55} M_{66} M_{12} M_{23} M_{34} M_{45} M_{56} M_{36}]^T$ .

TABLE IV  
COMPARISONS OF DIFFERENT DEEP NEURAL NETWORK TECHNIQUES

Deep Neural Network Techniques	Neural Network Structure	Model Training Complexity	Time-Domain Modeling with Memory Effect	Image-Like Inputs
Hybrid deep MLP [91]	Fully-Connected	Simple	No	Inefficient
Local-Feedback Deep RNN [133]	Fully-Connected with Feedback	Complex	Yes	Medium
Augmented CNN [172]	Partially-Connected	Medium	No	Efficient

The model inputs are  $S_{11}$  in dB at 41 frequency samples, i.e.,  $\bar{x} = [\text{dB}(S_{11}(f_1)) \text{dB}(S_{11}(f_2)) \cdots \text{dB}(S_{11}(f_{41}))]^T$ . In this example, each hidden layer contains 200 hidden neurons. After training with the three-stage deep learning algorithm, the structure of the hybrid deep MLP model for this sixth-order example is shown in Fig. 9. In the developed deep neural network model, eight hidden layers employ sigmoid functions, and four hidden layers employ smooth ReLUs. The training and test errors of the developed model are 1.31% and 1.79%, respectively [91].

After the deep MLP parameter-extraction model is developed, a detuned filter is used to examine the developed model. Fig. 10 shows the comparisons between the desired S-parameters and the S-parameters calculated from the extracted coupling matrix (i.e., outputs of the hybrid deep MLP model) for the detuned filter.

## VII. APPLICATIONS OF NEURAL NETWORKS IN MICROWAVE CAD

### A. EM Analysis and Parametric Modeling

EM parametric modeling is to provide EM behavior with variable values of physical/geometrical parameters of an EM structure. Such a model is important in EM-based design. The use of neural networks can help avoid repetitive EM simulations by learning the relationship between EM response and the varying values of geometrical parameters in advance. Many research activities have been devoted to such applications, for example, high-speed interconnects [93], CPW circuit components [94], practical multilayered shielded microwave circuits [124], spiral inductors [45], [95], EM interference (EMI) estimation [96], internally decomposed EM structure [97], metasurfaces [175], differential via holes [46], and couplers [98].

Bandler *et al.* [37] have presented a modeling method of microwave circuits using a combined ANN and space mapping (SM) technique. The SM concept combines the computational efficiency of coarse models (typically empirical functions/equivalent circuit models) with the accuracy of fine models (typically EM/physics models). Besides the SM technique, the RBF-NN models have been presented by

Soliman *et al.* [125] for the efficient filling of the coupling matrix of the method of moments (MoM). Ding *et al.* [99] have introduced the ANN approach to EM-based modeling of passive components in both frequency and time domains.

### B. Behavioral Modeling of Nonlinear Circuits

The behavioral modeling for nonlinear circuits is an interesting problem where the circuit responses are dependent on signal amplitude and terminations as well as frequencies. Simple analytical formulas often do not exist. Therefore, neural network becomes a good choice for this task. Various ANN techniques have been developed for behavioral modeling of nonlinear circuits. Novel neural network structures have been introduced to represent the dynamic relationships between nonlinear circuit inputs and outputs, for example, RNN [132], [134], [135], DNN [130], nonlinear model-order reduction method [107], neural network-based bidirectional and dispersive behavior models [25], real-valued TDNN (RVTDNN) [137], adjoint neural networks [176], envelope-domain time-series model [177], and the Wiener-type DNN [131].

Behavioral modeling of power amplifiers has been an interesting topic in the ANN modeling area. Various ANN modeling and design methods for power amplifiers have been introduced, e.g., [89], [90], [108], [122], [129], [138], [139], [142], [178], [179], [180], [181]. Related to power amplifier design, ANN has also been used in digital predistortion, e.g., [182], [183], [184], [185], [186].

### C. Modeling of Microwave Transistors

Modeling of microwave transistors is an important topic, especially when the behaviors of transistors become complex under high-power and high-frequency operations in modern wireless communications. The ANN method has become an attractive approach for modeling of microwave transistors, since it can provide smooth functions while providing the capability to fit complicated nonlinear bias/power dependence in devices. ANN methods have been used for modeling various types of microwave transistors, for example, field-effect transistors (FETs) [47], [109], [110], [187], heterojunction bipolar transistors (HBTs) [48], [49], [111], and high electron-mobility transistors (HEMTs) [112], [113], [114], [115], [136], [188].

Gallium nitride (GaN) devices, such as GaN HEMTs, are important for next-generation wireless communication systems and microwave power device applications. However, the GaN HEMT devices exhibit strong thermal effects and trapping effects that would deteriorate the performance of the RF and microwave circuits. As a consequence, an accurate large-signal model for GaN HEMT devices capable of characterizing the sophisticated thermal and trapping effects is the key to high-reliability RF and microwave circuit design. Examples of ANN-based GaN HEMT modeling are the ANN-based electrothermal model [106], [189], [190], [191], the ANN-based large-signal model [192], [193], the hybrid modeling approach using ANN [194], the SM technique using decomposed mappings [195], and the ANN-based consistent gate charge model [196].

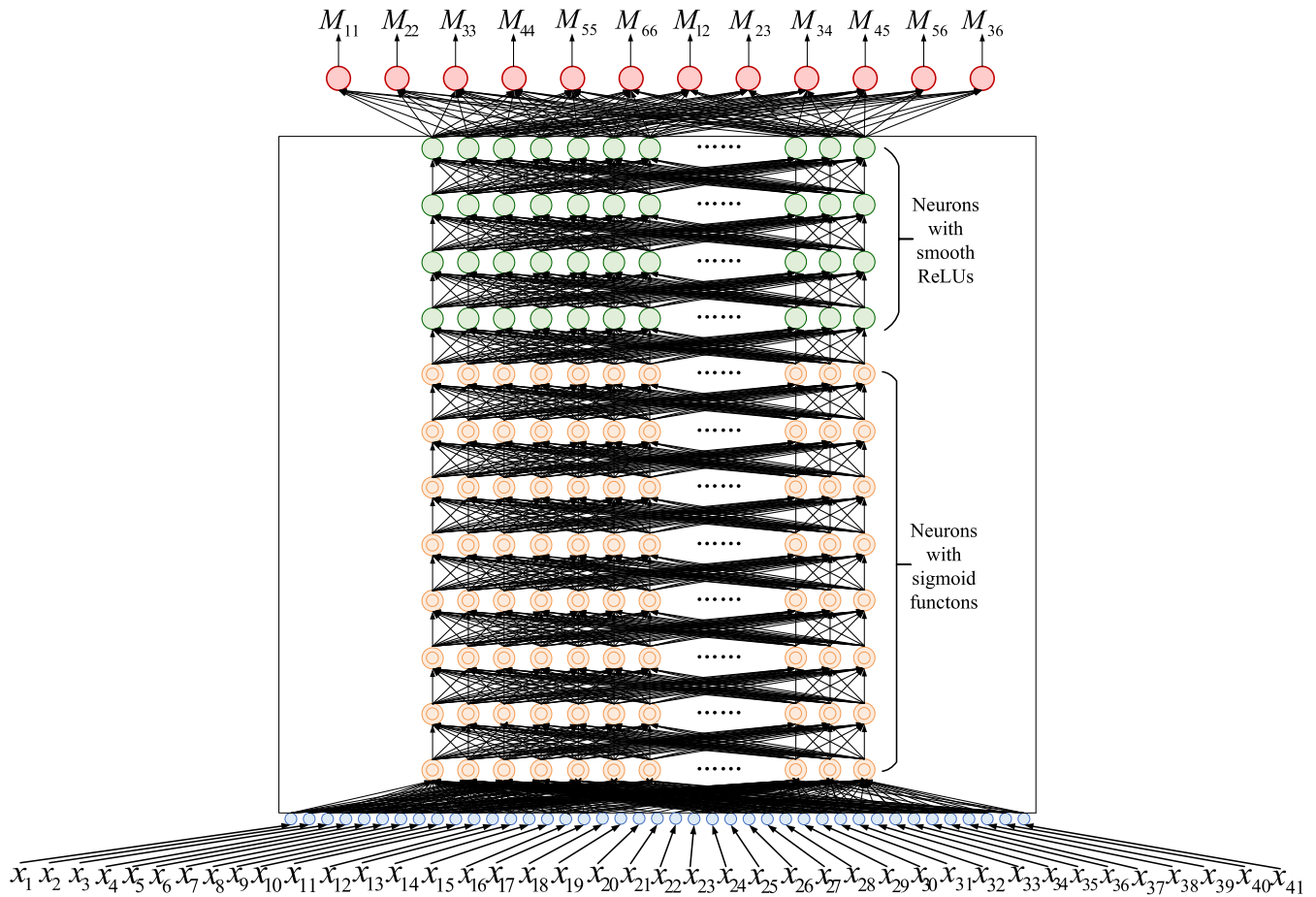


Fig. 9. Structure of the deep neural network high-dimensional parameter-extraction model for the six-order filter [91]. The 41 input neurons represent the dB value of  $S_{11}$  over 41 frequency points.

#### D. Filter Modeling and Design

Filter design is one of the most fundamental parts of microwave CAD. However, the full-wave EM simulation-based approach to filter design is computationally expensive. Various researchers have exploited the ANN approach to help filter modeling and design. For example, homotopy optimization based on the ANN model overcomes the problem that the initial value is not close enough to the optimal solution [197]. An efficient hybrid sampling method for ANN-based microwave component modeling and optimization can achieve an accurate ANN model with limited training data [116]. An efficient ELM with transfer functions improves the learning speed and accuracy [127], [128]. A high-dimensional modeling problem has been addressed in [50] and [198]. A novel deep-q-network (DQN)-based fine-tuning approach is developed to reduce the computation time even with different initial values [143]. Specific KBNN-based approaches are developed for different filter applications, for example, high-power dielectric resonator filters [199], narrow-bandwidth four-pole Ku-band bandpass filters [200], and even multiplexers [201].

#### E. VLSI Interconnects

Neural network techniques have also been introduced to very large-scale integration (VLSI) interconnects.

For example, Veluswami *et al.* [93] have discussed a neural network-based approach to the EM simulation and optimization of interconnects in high-speed VLSI circuits. Ding *et al.* [99] have applied ANN to EM-based signal integrity analysis and optimization of a high-speed VLSI interconnect circuit with embedded passive terminations and a nonlinear buffer. Ilumoka [202] has applied the modular ANN model consisting of groups of networks competing to learn different aspects of a problem to complex interactions of 3-D interconnect. To alleviate the time-consuming iterations of the full-wave simulations procedure, Hsu *et al.* [203] have utilized the neural network-based approach to efficient design of reflectionless via structures with single-ended interconnections. The deep neural network approach has been applied by Lu *et al.* [204] to high-speed channel modeling for signal integrity analysis, which saves complex circuit simulations and substantial domain knowledge altogether. Chiu *et al.* [205] have adopted ANN optimization to facilitate the design of signal integrity. Swaminathan *et al.* [206] have discussed the neural network applications for signal and power integrity in packaging. To overcome the problem that the reliability simulation process takes too much time and resources, Tian *et al.* [207] have introduced a neural network modeling method to analyze ultralarge-scale integration (ULSI) interconnect reliability.

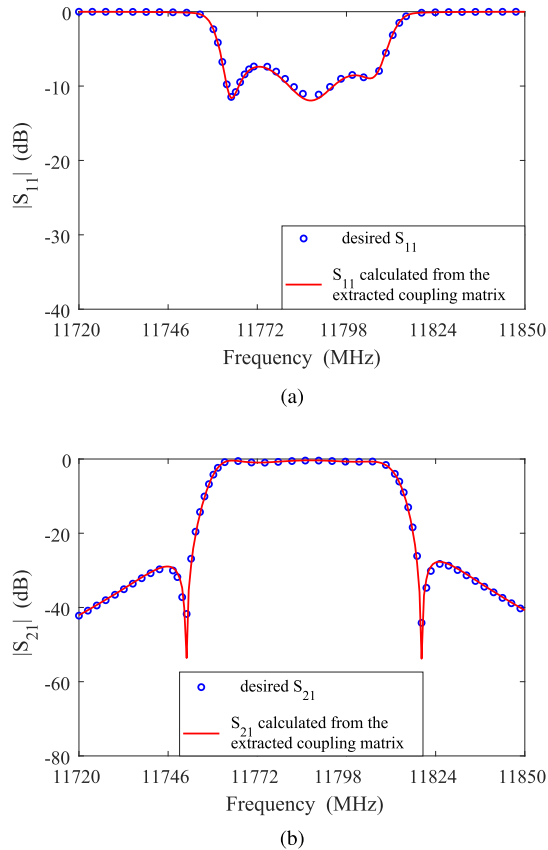


Fig. 10. Comparison between the desired S-parameters and the S-parameters calculated from the extracted coupling matrix (i.e., outputs of the deep neural network model) for the detuned sixth-order filter. (a) Return loss  $S_{11}$ . (b) Insertion loss  $S_{21}$  [91].

#### F. Behavioral Modeling Related to Oscillators

Behavioral modeling of oscillators is an important application of neural network techniques. For example, Yu *et al.* [117], [118] have developed a modeling method for voltage-controlled oscillators (VCOs) using augmented neural networks to improve computational efficiency. FFNNs are used to predict the oscillatory output waveforms, and RNNs are used to capture the nonlinear dynamic current–voltage relation at the output port. Sen *et al.* [119] have reported a neural-network-based parasitic modeling and extraction verification for RF/millimeter-wave integrated circuit design. They demonstrate the verification approach using automatically generated passive test structures and ring oscillators. This method helps improve the extraction–verification efficiency and accuracy.

#### G. Transmitter and Receiver Modeling and Design

Transmitter and receiver modeling and design using neural networks have gained attention in recent years. For example, Yin *et al.* [208] have discussed the behavioral modeling of millimeter-wave beamforming transmitters. They have utilized concurrent dynamic configurations with heterogeneous neural networks and achieved good performance on both training and validation sets, which provides a promising high-accuracy OTA model for mmW beamforming transmitters with concurrent configurations. Zhang *et al.* [209] have presented to use a RVTDDNN for the calibration of multiport wideband receivers.

They have considered all the system impairments, including the non-ideal six-port correlator, the nonlinearity of the diodes, and the memory effects under the excitation of the wideband signals. The performance achieved using the method in [209] for most commonly used communication signals is better than the previous methods.

#### H. Neural Network-Based Multiphysics Modeling

Multiphysics simulation of microwave components involves the simultaneous solutions of EM and other physics domains [210]. Multiphysics simulation can provide more accurate results but is more time-consuming than pure EM simulation. The ANN technique has also been introduced to the multiphysics modeling and design of microwave components. For example, neural network-based modeling approaches have been introduced to improve the efficiency of multiphysics modeling with limited training samples [100], [101]. The multiphysics optimization technique has also been developed using KBNNs to accelerate the multiphysics design process [102], [103], [104].

#### I. Further Applications of ANN in Microwave CAD

There are further applications of ANN in microwave CAD. For example, Kabir *et al.* [120] have discussed a neural network method for RFIC inductors to make the large space model efficient. The neural network model is used to compute coefficient values for the empirical equation, which produces the values of lumped elements of the double Pi network. An ANN-based modeling approach for the design of RF microelectromechanical system (MEMS) switches has been presented by Lee and Filipovic [105] to improve the modeling efficiency while maintaining the modeling accuracy. Gall *et al.* [211] have applied neural networks to spectrum-based direction-of-arrival estimation for automotive radar to achieve super-resolution performance and the capability of estimating. Jeong *et al.* [212] have discussed the implementation of a neural network-based machine learning strategy for real-time range-adaptive automatic impedance matching of wireless power transfer (WPT) applications. This method can effectively predict the optimal parameters of a tunable matching network and select the range of adaptive transmitting coils. Viveros-Wacher *et al.* [121] have applied an analog gross fault diagnosis method based on ANNs and constrained parameter extraction to analog fault diagnosis. Wei *et al.* [213] have introduced self-attention bidirectional LSTM (Bi-LSTM) networks for radar signal modulation recognition. The developed model has better recognition performance, especially at low SNRs with much lower computational complexity. Klinefelter and Nanzer [214] have applied the ANN-based approach to design millimeter-wave interferometric radar for automotive velocity sensing to achieve better estimation than the Bayesian estimator.

## VIII. CONCLUSION AND FUTURE DIRECTIONS

This article has reviewed the state of the art of the ANN techniques for microwave CAD. Fundamental concepts and



advanced techniques of ANN structure and training have been introduced. The KBNN has been presented for improving the accuracy and reliability of modeling and design optimization. Various ANN techniques, such as neuro-TF, neural network inverse modeling, and hybrid deep neural network technique, have been discussed. Further applications of ANN in microwave CAD have also been reviewed in this article. The literature review in this article covers the samples of the existing works on microwave ANN. However, our review is not exhaustive, and there are more papers existing in the literature, which are also valuable contributions to microwave ANN.

There are tremendous opportunities for further innovations and applications of ANN for microwave modeling and design, from advanced microwave-oriented ANN structures and training algorithms to novel applications. New methods will be needed for ANN models capable of wider input range, higher input dimensions, and, at the same time, allowing further reductions in the need of large amount of training data. Since the quality of ANN inherently depends on the quality and adequacy of training data, yet the expense of data generation in microwave problems is typically high, further advances in microwave-oriented ANN structures, data sampling, and training methods will remain important. Further advances in knowledge-based modeling methods continue to be another important direction, exploring new ways to incorporate different microwave knowledge into various intermediate parts of neural structures and training, and more generic methods to combine mapping of different subspaces of knowledge and neural network models. Incorporation of EM and multiphysics internal formulations into neural-based modeling process may lead to new solutions to conquer the expense of EM/multiphysics-based parametric modeling. Design is inherently an inverse problem to simulation. Subsequently, exploration of inverse modeling for fast and direct design solutions continues to be an attractive direction. Coupled with this design-oriented direction, there are many design steps that presently heavily require human experience and judgment. The exploration of machine learning algorithms to mimic the process of knowledge abstraction and knowledge utilization to accelerate the design process continues to be an open and strategic direction.

#### ACKNOWLEDGMENT

The author Qi-Jun Zhang would like to thank Prof. Jianguo Ma, Editor-in-Chief of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES, for the initial invitation to write this review paper.

#### REFERENCES

- [1] Q.-J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*. Norwood, MA, USA: Artech House, 2000.
- [2] Q.-J. Zhang, K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for RF and microwave design—From theory to practice," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 4, pp. 1339–1350, Apr. 2003.
- [3] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 420–435, Jan. 2004.
- [4] V. Rizzoli, A. Costanzo, D. Masotti, A. Lipparini, and F. Mastrì, "Computer-aided optimization of nonlinear microwave circuits with the aid of electromagnetic simulation," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 362–377, Jan. 2004.
- [5] M. B. Steer, J. W. Bandler, and C. M. Snowden, "Computer-aided design of RF and microwave circuits and systems," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 3, pp. 996–1005, Mar. 2002.
- [6] M. Vai and S. Prasad, "Automatic impedance matching with a neural network," *IEEE Microw. Guided Wave Lett.*, vol. 3, no. 10, pp. 353–354, Oct. 1993.
- [7] T.-S. Horng, C.-C. Wang, and N. G. Alexopoulos, "Microstrip circuit design using neural networks," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1993, pp. 413–416.
- [8] P. Watson and K. C. Gupta, "EM-ANN models for via interconnects in microstrip circuits," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1996, pp. 1819–1822.
- [9] P. M. Watson and K. C. Gupta, "EM-ANN models for microstrip vias and interconnects in dataset circuits," *IEEE Trans. Microw. Theory Techn.*, vol. 44, no. 12, pp. 2495–2503, Dec. 1996.
- [10] G. L. Creech, B. Paul, C. Lesniak, T. Jenkins, R. Lee, and M. Calcaterra, "Artificial neural networks for accurate microwave CAD applications," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1996, pp. 733–736.
- [11] M. Vai and S. Prasad, "Microwave circuit analysis and design by a massively distributed computing network," *IEEE Trans. Microw. Theory Techn.*, vol. 43, no. 5, pp. 1087–1094, May 1995.
- [12] A. H. Zaabab, Q. J. Zhang, and M. Nakhla, "Analysis and optimization of microwave circuits and devices using neural network models," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 1994, pp. 393–396.
- [13] A. H. Zaabab, Q.-J. Zhang, and M. Nakhla, "A neural network modeling approach to circuit optimization and statistical design," *IEEE Trans. Microw. Theory Techn.*, vol. 43, no. 6, pp. 1349–1358, Jun. 1995.
- [14] P. Burrascano, S. Fiori, and M. Mongiardo, "A review of artificial neural networks applications in microwave computer-aided design," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 158–174, May 1999.
- [15] F. Wang, V. K. Devabhaktuni, C. Xi, and Q.-J. Zhang, "Neural network structures and training algorithms for RF and microwave applications," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 216–240, May 1999.
- [16] V. K. Devabhaktuni, C. Xi, F. Wang, and Q.-J. Zhang, "Robust training of microwave neural models," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 109–124, Jan. 2002.
- [17] P. M. Watson, C. Cho, and K. C. Gupta, "Electromagnetic-artificial neural network model for synthesis of physical dimensions for multi-layer asymmetric coupled transmission structures," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 175–186, May 1999.
- [18] S. Selleri, S. Manetti, and G. Pelosi, "Neural network applications in microwave device design," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 90–97, Jan. 2002.
- [19] R. K. Mishra, "An overview of neural network methods in computational electromagnetics," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 98–108, Jan. 2002.
- [20] G. Fedi, A. Gaggelli, S. Manetti, and G. Pelosi, "Direct-coupled cavity filters design using a hybrid feedforward neural network-finite elements procedure," *Int. J. RF Microw. Comp.-Aided Eng.*, vol. 9, no. 3, pp. 287–296, May 1999.
- [21] S. Bila *et al.*, "An accurate wavelet neural-network-based model for electromagnetic optimization of microwave circuits," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 297–306, May 1999.
- [22] J. W. Bandler, J. E. Rayas-Sánchez, and Q.-J. Zhang, "Yield-driven electromagnetic optimization via space mapping-based neuromodels," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, Aug. 1, pp. 79–89, Jan. 2002.
- [23] M. Vai and S. Prasad, "Neural networks in microwave circuit design—beyond black-box models," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 187–197, May 1999.
- [24] Y. Harkouss, J. Rousset, H. Chehade, E. Ngoya, D. Barataud, and J. P. Teysier, "The use of artificial neural networks in nonlinear microwave devices and circuits modeling: An application to telecommunication system design," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 198–215, May 1999.
- [25] V. Rizzoli, A. Neri, D. Masotti, and A. Lipparini, "A new family of neural network-based bidirectional and dispersive behavioral models for nonlinear RF/microwave subsystems," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 51–70, Jan. 2002.

- [26] C. Li, S. He, X. Liao, and J. Yu, "Using recurrent neural network for adaptive predistortion linearization of RF amplifiers," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 125–130, Jan. 2002.
- [27] J. A. García *et al.*, "Modeling MESFETs and HEMTs intermodulation distortion behavior using a generalized radial basis function network," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 261–276, May 1999.
- [28] P. M. Watson, K. C. Gupta, and R. L. Mahajan, "Applications of knowledge-based artificial neural network modeling to microwave components," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 254–260, May 1999.
- [29] G. L. Creech and J. M. Zurada, "Neural network modeling of GaAs IC material and MESFET device characteristics," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 241–253, May 1999.
- [30] Ş. Sağıroğlu, K. Gıney, and M. Erler, "Calculation of bandwidth for electrically thin and thick rectangular microstrip antennas with the use of multilayered perceptrons," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 9, no. 3, pp. 277–286, May 1999.
- [31] K. Guney, S. Sağıroğlu, and M. Erler, "Generalized neural method to determine resonant frequencies of various microstrip antennas," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 131–139, Jan. 2002.
- [32] S. Goasguen and S. M. El-Ghazaly, "A coupled FDTD-artificial neural network technique for large-signal analysis of microwave circuits," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 25–36, Jan. 2002.
- [33] F. Giannini, G. Leuzzi, G. Orenco, and M. Albertini, "Small-signal and large-signal modeling of active devices using CAD-optimized neural networks," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 71–78, Jan. 2002.
- [34] G. Pan, P. Piel, and B. Gilbert, "A hybrid integral equation and neural network approach for fast extraction of frequency dependent parameters of multiconductor transmission lines," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 37–50, Jan. 2002.
- [35] J. A. Jargon, K. C. Gupta, and D. C. DeGroot, "Applications of artificial neural networks to RF and microwave measurements," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 12, no. 1, pp. 3–24, Jan. 2002.
- [36] F. Wang and Q. J. Zhang, "Knowledge based neural models for microwave design," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1997, pp. 627–630.
- [37] J. W. Bandler, M. A. Ismail, J. E. Rayas-Sánchez, and Q.-J. Zhang, "Neuromodeling of microwave circuits exploiting space-mapping technology," *IEEE Trans. Microw. Theory Techn.*, vol. 47, no. 12, pp. 2417–2427, Dec. 1999.
- [38] F. Dandurand and D. A. Lowther, "Electromagnetic device performance identification using knowledge based neural networks," *IEEE Trans. Magn.*, vol. 35, no. 3, pp. 1817–1820, May 1999.
- [39] R. Varma and J. Ghosh, "Modeling of a compact triple band PIFA using knowledge based neural network," in *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2016, pp. 2190–2192.
- [40] T. V. S. Divakar and D. C. Panda, "Resonant frequency calculation of circular ring microstrip patch antenna using knowledge based neural network," in *Proc. Int. Conf. Electron. Commun. Syst. (ICECS)*, Feb. 2014, pp. 1–3.
- [41] R. Kumar *et al.*, "Knowledge-based neural networks for fast design space exploration of hybrid copper-graphene on-chip interconnect networks," *IEEE Trans. Electromagn. Compat.*, vol. 64, no. 1, pp. 182–195, Feb. 2022.
- [42] L. Zhang, J. Xu, M. C. E. Yagoub, R. Ding, and Q. Zhang, "Neuro-space mapping technique for nonlinear device modeling and large signal simulation," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2003, pp. 173–176.
- [43] L. Zhang, Q. J. Zhang, and J. Wood, "Statistical neuro-space mapping technique for large-signal modeling of nonlinear devices," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 11, pp. 2453–2467, Nov. 2008.
- [44] W. Na, F. Feng, C. Zhang, and Q.-J. Zhang, "A unified automated parametric modeling algorithm using knowledge-based neural network and  $l_1$  optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 3, pp. 729–745, Dec. 2017.
- [45] Y. Cao and G. Wang, "A wideband and scalable model of spiral inductors using space-mapping neural network," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 12, pp. 2473–2480, Dec. 2007.
- [46] Y. Cao, L. Simonovich, and Q. J. Zhang, "A broadband and parametric model of differential via holes using space-mapping neural network," *IEEE Microw. Wireless Compon. Lett.*, vol. 19, no. 9, pp. 533–535, Sep. 2009.
- [47] D. Gorissen, L. Zhang, Q.-J. Zhang, and T. Dhaene, "Evolutionary neuro-space mapping technique for modeling of nonlinear microwave devices," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 2, pp. 213–229, Feb. 2011.
- [48] L. Zhang, J. Xu, M. C. E. Yagoub, R. Ding, and Q.-J. Zhang, "Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 9, pp. 2752–2767, Sep. 2005.
- [49] M. Iwamoto, J. Xu, W. Zhou, and D. E. Root, "Knowledge-based neural network (KBNN) modeling of HBT junction temperature and thermal resistance from electrical measurements," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 1069–1071.
- [50] H. Kabir, Y. Wang, M. Yu, and Q.-J. Zhang, "High-dimensional neural-network technique and applications to microwave filter modeling," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 145–156, Jan. 2010.
- [51] V. K. Devabhaktuni, B. Chattaraj, M. C. E. Yagoub, and Q.-J. Zhang, "Advanced microwave modeling framework exploiting automatic model generation, knowledge neural networks, and space mapping," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 7, pp. 1822–1833, Jul. 2003.
- [52] J. E. Rayas-Sánchez and V. Gutiérrez-Ayala, "EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 12, pp. 4528–4537, Dec. 2006.
- [53] H. Kabir, L. Zhang, M. Yu, P. H. Aaen, J. Wood, and Q.-J. Zhang, "Smart modeling of microwave devices," *IEEE Microw. Mag.*, vol. 11, no. 3, pp. 105–118, May 2010.
- [54] Y. Cao, G. Wang, and Q.-J. Zhang, "A new training approach for parametric modeling of microwave passive components using combined neural networks and transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 11, pp. 2727–2742, Nov. 2009.
- [55] F. Feng, C. Zhang, J. Ma, and Q.-J. Zhang, "Parametric modeling of EM behavior of microwave components using combined neural networks and pole-residue-based transfer functions," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 1, pp. 60–77, Jan. 2016.
- [56] J. Zhang, F. Feng, W. Zhang, J. Jin, J. Ma, and Q.-J. Zhang, "A novel training approach for parametric modeling of microwave passive components using Padé via Lanczos and EM sensitivities," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 6, pp. 2215–2233, Jun. 2020.
- [57] F. Feng and Q. J. Zhang, "Parametric modeling using sensitivity-based adjoint neuro-transfer functions for microwave passive components," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2015, pp. 1–3.
- [58] F. Feng, V.-M.-R. Gongal-Reddy, C. Zhang, J. Ma, and Q. J. Zhang, "Parametric modeling of microwave components using adjoint neural networks and pole-residue transfer functions with EM sensitivity analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 6, pp. 1955–1975, Jun. 2017.
- [59] J. Zhang, J. Chen, Q. Guo, W. Liu, F. Feng, and Q.-J. Zhang, "Parameterized modeling incorporating MOR-based rational transfer functions with neural networks for microwave components," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 5, pp. 379–382, May 2022.
- [60] Z. Guo, J. Gao, Y. Cao, and Q.-J. Zhang, "Passivity enforcement for passive component modeling subject to variations of geometrical parameters using neural networks," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2012, pp. 1–3.
- [61] H. Wang, L. Sun, J. Liu, H. Zou, Z. Yu, and J. Gao, "Transfer function analysis and broadband scalable model for on-chip spiral inductors," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 7, pp. 1696–1708, Jul. 2011.
- [62] D. Spina, D. De Jonghe, D. Deschrijver, G. Gielen, L. Knockaert, and T. Dhaene, "Stochastic macromodeling of nonlinear systems via polynomial chaos expansion and transfer function trajectories," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 7, pp. 1454–1460, Jul. 2014.
- [63] Y. Eo and W. R. Eisenstadt, "Generalized coupled interconnect transfer function and high-speed signal simulations," *IEEE Trans. Microw. Theory Techn.*, vol. 43, no. 5, pp. 1115–1121, May 1995.
- [64] D.-C. Chang and C.-W. Hsue, "Design and implementation of filters using transfer functions in the Z domain," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 5, pp. 979–985, May 2001.

- [65] A. Zadehgo, "An impedance transfer function formulation for reduced-order macromodels of subgridded regions in FDTD," *IEEE Trans. Antennas Propag.*, vol. 65, no. 1, pp. 401–404, Jan. 2017.
- [66] G. J. Burke, E. K. Miller, S. Chakrabarti, and K. Demarest, "Using model-based parameter estimation to increase the efficiency of computing electromagnetic transfer functions," *IEEE Trans. Magn.*, vol. 25, no. 4, pp. 2807–2809, Jul. 1989.
- [67] H. Tiberg, A. Bergqvist, L. Kvarnsjo, and G. Engdahl, "A method of computing electro-mechanical transfer functions for magnetostrictive functional units," *IEEE Trans. Magn.*, vol. 28, no. 5, pp. 2808–2810, Sep. 1992.
- [68] F. Feng, C. Zhang, W. Na, J. Zhang, W. Zhang, and Q. J. Zhang, "Adaptive feature zero assisted surrogate-based EM optimization for microwave filter design," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 1, pp. 2–4, Jan. 2019.
- [69] 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.
- [70] F. Feng, W. Na, W. Liu, S. Yan, L. Zhu, and Q.-J. Zhang, "Parallel gradient-based EM optimization for microwave components using adjoint-sensitivity-based neuro-transfer function surrogate," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 9, pp. 3606–3620, Sep. 2020.
- [71] J. Zhang, F. Feng, and Q. J. Zhang, "Rapid yield estimation of microwave passive components using model-order reduction based neuro-transfer function models," *IEEE Microw. Wireless Comp. Lett.*, vol. 31, no. 4, pp. 333–336, Apr. 2021.
- [72] J. Zhang, F. Feng, J. Jin, W. Zhang, Z. Zhao, and Q.-J. Zhang, "Adaptively weighted yield-driven EM optimization incorporating neuro-transfer function surrogate with applications to microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 1, pp. 518–528, Jan. 2021.
- [73] H. Acikgoz, B. Jannier, Y. L. Bihan, O. Dubrunfaut, O. Meyer, and L. Pichon, "Direct and inverse modeling of a microwave sensor determining the proportion of fluids in a pipeline," *IEEE Trans. Magn.*, vol. 45, no. 3, pp. 1510–1513, Mar. 2009.
- [74] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 1819–1825, Mar. 2019.
- [75] L. Yuan, X.-S. Yang, C. Wang, and B.-Z. Wang, "Inverse design of metasurface using combined neural network and transfer function," in *Proc. IEEE Int. Symp. Antennas Propag. North Amer. Radio Sci. Meeting*, Jul. 2020, pp. 2027–2028.
- [76] G. Gosal, E. Almajali, D. McNamara, and M. Yagoub, "Transmitarray antenna design using forward and inverse neural network modeling," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 1483–1486, 2016.
- [77] Z. Lin *et al.*, "Low-frequency data prediction with iterative learning for highly nonlinear inverse scattering problems," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 10, pp. 4366–4376, Oct. 2021.
- [78] 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.
- [79] 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.*, early access, Apr. 8, 2022, doi: [10.1109/TMTT.2022.3161928](https://doi.org/10.1109/TMTT.2022.3161928).
- [80] H. Kabir, Y. Wang, M. Yu, and Q.-J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 4, pp. 867–879, Apr. 2008.
- [81] L. Yuan, X.-S. Yang, C. Wang, and B.-Z. Wang, "Multibranch artificial neural network modeling for inverse estimation of antenna array directivity," *IEEE Trans. Antennas Propag.*, vol. 68, no. 6, pp. 4417–4427, Jun. 2020.
- [82] Y. Sato, K. Kawano, and H. Igarashi, "Direct inverse modeling for electromagnetic components using Gaussian kernel regression," *IEEE Trans. Magn.*, vol. 58, no. 5, pp. 1–8, May 2022.
- [83] 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.
- [84] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [85] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Trans. Antennas Propag.*, vol. 68, no. 7, pp. 5626–5635, Jul. 2020.
- [86] A. Tamminen, S.-V. Palli, J. Ala-Laurinaho, and Z. D. Taylor, "Millimeter- and submillimeter-wave imaging through dispersive hologram and deep neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 6, pp. 3281–3290, Jun. 2022.
- [87] G. Cheng, M. Wang, W. Zhang, and X. Liu, "Advanced deep neural network technique for microwave parametric modeling," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2021, pp. 1–3.
- [88] G. W. Cottrell, "New life for neural networks," *Science*, vol. 313, no. 5786, pp. 454–455, Jul. 2006.
- [89] Z. Liu, X. Hu, T. Liu, X. Li, W. Wang, and F. M. Ghannouchi, "Attention-based deep neural network behavioral model for wideband wireless power amplifiers," *IEEE Microw. Wireless Compon. Lett.*, vol. 30, no. 1, pp. 82–85, Jan. 2020.
- [90] S. Zhang *et al.*, "Deep neural network behavioral modeling based on transfer learning for broadband wireless power amplifier," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 7, pp. 917–920, Jul. 2021.
- [91] J. Jin, C. Zhang, F. Feng, W. Na, J. Ma, and Q.-J. Zhang, "Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 10, pp. 4140–4155, Oct. 2019.
- [92] S. A. Sadrossadat, Y. Cao, and Q.-J. Zhang, "Parametric modeling of microwave passive components using sensitivity-analysis-based adjoint neural-network technique," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 1733–1747, May 2013.
- [93] A. Veluswami, M. S. Nakhla, and Q.-J. Zhang, "The application of neural networks to EM-based simulation and optimization of interconnects in high-speed VLSI circuits," *IEEE Trans. Microw. Theory Techn.*, vol. 45, no. 5, pp. 712–723, May 1997.
- [94] P. M. Watson and K. C. Gupta, "Design and optimization of CPW circuits using EM-ANN models for CPW components," *IEEE Trans. Microw. Theory Techn.*, vol. 45, no. 12, pp. 2515–2523, Dec. 1997.
- [95] S. K. Mandal, S. Sural, and A. Patra, "ANN- and PSO-based synthesis of on-chip spiral inductors for RF ICs," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 1, pp. 188–192, Jan. 2008.
- [96] Y.-F. Shu, X.-C. Wei, J. Fan, R. Yang, and Y.-B. Yang, "An equivalent dipole model hybrid with artificial neural network for electromagnetic interference prediction," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 5, pp. 1790–1797, May 2019.
- [97] S. Liao, H. Kabir, Y. Cao, J. Xu, Q.-J. Zhang, and J. Ma, "Neural-network modeling for 3-D substructures based on spatial EM-field coupling in finite-element method," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 1, pp. 21–38, Jan. 2011.
- [98] A. Liu, M. Leng, G. Pan, and M. Yu, "Automatic coupler design using data-driven-based modular neural network," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 12, pp. 1259–1262, Dec. 2021.
- [99] X. Ding *et al.*, "Neural-network approaches to electromagnetic-based modeling of passive components and their applications to high-frequency and high-speed nonlinear circuit optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 436–449, Jan. 2004.
- [100] 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.
- [101] W. Zhang, F. Feng, J. Zhang, S. Zhang, V.-M.-R. Gongal-Reddy, and Q. J. Zhang, "Advanced parametric modeling using neuro-transfer function for EM based multiphysics analysis of microwave passive components," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2016, pp. 1–3.
- [102] W. Zhang, F. Feng, S. Yan, W. Na, J. Ma, and Q.-J. Zhang, "EM-centric multiphysics optimization of microwave components using parallel computational approach," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 2, pp. 479–489, Feb. 2020.
- [103] W. Zhang, F. Feng, J. Jin, and Q.-J. Zhang, "Parallel multiphysics optimization for microwave devices exploiting neural network surrogate," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 4, pp. 341–344, Apr. 2021.
- [104] 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.
- [105] Y. Lee and D. S. Filipovic, "ANN based electromagnetic models for the design of RF MEMS switches," *IEEE Microw. Wireless Compon. Lett.*, vol. 15, no. 11, pp. 823–825, Nov. 2005.

- [106] A.-D. Huang, Z. Zhong, W. Wu, and Y.-X. Guo, "An artificial neural network-based electrothermal model for GaN HEMTs with dynamic trapping effects consideration," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 8, pp. 2519–2528, Aug. 2016.
- [107] J. Wood, D. E. Root, and N. B. Tuffillaro, "A behavioral modeling approach to nonlinear model-order reduction for RF/microwave ICs and systems," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 9, pp. 2274–2284, Sep. 2004.
- [108] J. Louro *et al.*, "New transistor behavioral model formulation suitable for Doherty PA design," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 4, pp. 2138–2147, Apr. 2021.
- [109] Y. Ko *et al.*, "Artificial neural network model of SOS-MOSFETs based on dynamic large-signal measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 3, pp. 491–501, Mar. 2014.
- [110] S. Zhang, J. Xu, Q.-J. Zhang, and D. E. Root, "Parallel matrix neural network training on cluster systems for dynamic FET modeling from large datasets," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2016, pp. 1–3.
- [111] H. Taher, D. Schreurs, and B. Nauwelaers, "Constitutive relations for nonlinear modeling of Si/SiGe HBTs using an ANN model," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 15, no. 2, pp. 203–209, Mar. 2005.
- [112] K. Shirakawa, M. Shimizu, N. Okubo, and Y. Daido, "Structural determination of multilayered large-signal neural-network HEMT model," *IEEE Trans. Microw. Theory Techn.*, vol. 46, no. 10, pp. 1367–1375, Oct. 1998.
- [113] Y. Long, Z. Zhong, and Y.-X. Guo, "A novel 4-D artificial-neural-network-based hybrid large-signal model of GaAs pHEMTs," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 6, pp. 1752–1762, Jun. 2016.
- [114] L. Zhu, Q. Zhang, K. Liu, Y. Ma, B. Peng, and S. Yan, "A novel dynamic neuro-space mapping approach for nonlinear microwave device modeling," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 2, pp. 131–133, Feb. 2016.
- [115] J. Gao, L. Shen, and D. Luo, "High frequency HEMT modeling using artificial neural network technique," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2015, pp. 1–3.
- [116] Z. Zhang, Q. S. Cheng, H. Chen, and F. Jiang, "An efficient hybrid sampling method for neural network-based microwave component modeling and optimization," *IEEE Microw. Wireless Compon. Lett.*, vol. 30, no. 7, pp. 625–628, Jul. 2020.
- [117] H. Yu, H. Chalamalasetty, and M. Swaminathan, "Modeling of voltage-controlled oscillators including I/O behavior using augmented neural networks," *IEEE Access*, vol. 7, pp. 38973–38982, 2019.
- [118] H. Yu, M. Swaminathan, C. Ji, and D. White, "A nonlinear behavioral modeling approach for voltage-controlled oscillators using augmented neural networks," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2018, pp. 551–554.
- [119] P. Sen *et al.*, "Neural-network-based parasitic modeling and extraction verification for RF/millimeter-wave integrated circuit design," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 6, pp. 2604–2614, Jun. 2006.
- [120] H. Kabir, L. Zhang, and K. Kim, "Automatic parametric model development technique for RFIC inductors with large modeling space," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 551–554.
- [121] A. Viveros-Wacher, J. E. Rayas-Sanchez, and Z. Brito-Brito, "Analog gross fault identification in RF circuits using neural models and constrained parameter extraction," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2143–2150, Jun. 2019.
- [122] M. Isaksson, D. Wisell, and W. Rönnow, "Wide-band dynamic modeling of power amplifiers using radial-basis function neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 11, pp. 3422–3428, Nov. 2005.
- [123] L.-Y. Xiao, W. Shao, F.-L. Jin, B.-Z. Wang, W. T. Joines, and Q. H. Liu, "Semisupervised radial basis function neural network with an effective sampling strategy," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 4, pp. 1260–1269, Apr. 2020.
- [124] J. P. Garcia, F. Q. Pereira, D. C. Rebenague, J. L. G. Tornero, and A. A. Melcon, "A neural-network method for the analysis of multilayered shielded microwave circuits," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 1, pp. 309–320, Jan. 2006.
- [125] E. A. Soliman, M. H. Bakr, and N. K. Nikolova, "Neural networks-method of moments (NN-MoM) for the efficient filling of the coupling matrix," *IEEE Trans. Antennas Propag.*, vol. 52, no. 6, pp. 1521–1529, Jun. 2004.
- [126] J. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Comput.*, vol. 1, no. 2, pp. 281–294, Jun. 1989.
- [127] L.-Y. Xiao, W. Shao, T.-L. Liang, and B.-Z. Wang, "Efficient extreme learning machine with transfer functions for filter design," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 555–557.
- [128] L.-Y. Xiao, W. Shao, X. Ding, and B.-Z. Wang, "Dynamic adjustment kernel extreme learning machine for microwave component design," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 10, pp. 4452–4461, Oct. 2018.
- [129] C.-Y. Zhang, Y.-Y. Zhu, Q.-F. Cheng, H.-P. Fu, J.-G. Ma, and Q.-J. Zhang, "Extreme learning machine for the behavioral modeling of RF power amplifiers," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 558–561.
- [130] J. Xu, M. C. E. Yagoub, R. Ding, and Q.-J. Zhang, "Neural-based dynamic modeling of nonlinear microwave circuits," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 12, pp. 2769–2780, Dec. 2002.
- [131] W. Liu, W. Na, L. Zhu, J. Ma, and Q.-J. Zhang, "A Wiener-type dynamic neural network approach to the modeling of nonlinear microwave devices," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 6, pp. 2043–2062, Feb. 2017.
- [132] Y. Fang, M. C. E. Yagoub, F. Wang, and Q.-J. Zhang, "A new macro-modeling approach for nonlinear microwave circuits based on recurrent neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 48, no. 12, pp. 2335–2344, Dec. 2000.
- [133] M. Noohi, A. Mirvakili, and S. A. Sadrossadat, "Modeling and implementation of nonlinear boost converter using local feedback deep recurrent neural network for voltage balancing in energy harvesting applications," *Int. J. Circuit Theory Appl.*, vol. 49, no. 12, pp. 4231–4247, Sep. 2021.
- [134] B. Mutnury, M. Swaminathan, and J. P. Libous, "Macromodeling of nonlinear digital I/O drivers," *IEEE Trans. Adv. Packag.*, vol. 29, no. 1, pp. 102–113, Feb. 2006.
- [135] B. O'Brien, J. Dooley, and T. J. Brazil, "RF power amplifier behavioral modeling using a globally recurrent neural network," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2006, pp. 1089–1092.
- [136] Y. Cao, X. Chen, and G. Wang, "Dynamic behavioral modeling of nonlinear microwave devices using real-time recurrent neural network," *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1020–1026, May 2009.
- [137] T. Liu, S. Boumaiza, and F. M. Ghannouchi, "Dynamic behavioral modeling of 3G power amplifiers using real-valued time-delay neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 3, pp. 1025–1033, Mar. 2004.
- [138] J. Dooley, B. O'Brien, and T. J. Brazil, "Prior knowledge input neural network for microwave power amplifier modeling," in *Proc. Int. Workshop Integr. Nonlinear Microw. Millim.-Wave Circuits*, Jan. 2006, pp. 156–159.
- [139] Z. Zhao, W. Na, V.-M.-R. Gongal-Reddy, and Q. Zhang, "Multi-band behavioral modeling of power amplifier using carrier frequency-dependent time delay neural network model," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2017, pp. 82–84.
- [140] L. Kouhalvandi, O. Ceylan, and S. Ozoguz, "Automated deep neural learning-based optimization for high performance high power amplifier designs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 12, pp. 4420–4433, Dec. 2020.
- [141] M. A. Nielsen, *Neural Networks and Deep Learning*. Princeton, NJ, USA: Determination Press, 2015.
- [142] X. Yu, X. Hu, Z. Liu, C. Wang, W. Wang, and F. M. Ghannouchi, "A method to select optimal deep neural network model for power amplifiers," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 2, pp. 145–148, Feb. 2021.
- [143] M. Ohira, K. Takano, and Z. Ma, "A novel deep-Q-network-based fine-tuning approach for planar bandpass filter design," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 6, pp. 638–641, Jun. 2021.
- [144] R. Hongyo, Y. Egashira, T. M. Hone, and K. Yamaguchi, "Deep neural network-based digital predistorter for Doherty power amplifiers," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 2, pp. 146–148, Feb. 2019.
- [145] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2010, pp. 253–256.
- [146] Z. Wei and X. Chen, "Physics-inspired convolutional neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 9, pp. 6138–6148, Sep. 2019.

- [147] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
- [148] G. Pan, Y. Wu, M. Yu, L. Fu, and H. Li, "Inverse modeling for filters using a regularized deep neural network approach," *IEEE Microw. Wireless Compon. Lett.*, vol. 30, no. 5, pp. 457–460, Apr. 2020.
- [149] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [150] M. Moradi A., S. A. Sadrossadat, and V. Derhami, "Long short-term memory neural networks for modeling nonlinear electronic components," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 11, no. 5, pp. 840–847, May 2021.
- [151] O. W. Bhatti, N. Ambasana, and M. Swaminathan, "Design space and frequency extrapolation: Using neural networks," *IEEE Microw. Mag.*, vol. 22, no. 10, pp. 22–36, Oct. 2021.
- [152] S. Koziel, A. Pietrenko-Dabrowska, and U. Ullah, "Low-cost modeling of microwave components by means of two-stage inverse/forward surrogates and domain confinement," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 12, pp. 5189–5202, Dec. 2021.
- [153] M. D. Mckay, R. J. Beckman, and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, May 1979.
- [154] 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.
- [155] B. Pu, H. Kim, X.-D. Cai, B. Sen, C. Sui, and J. Fan, "Training set optimization in an artificial neural network constructed for high bandwidth interconnects design," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 6, pp. 2955–2964, Jun. 2022.
- [156] A. Saadallah and K. Morik, "Active sampling for learning interpretable surrogate machine learning models," in *Proc. IEEE 7th Int. Conf. Data Sci. Adv. Anal. (DSAA)*, Oct. 2020, pp. 264–272.
- [157] J. Shao, Q. Wang, and F. Liu, "Learning to sample: An active learning framework," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2019, pp. 538–547.
- [158] V. K. Devabhaktuni, M. C. E. Yagoub, and Q.-J. Zhang, "A robust algorithm for automatic development of neural-network models for microwave applications," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 12, pp. 2282–2291, Dec. 2001.
- [159] J. Cui, F. Feng, W. Na, and Q.-J. Zhang, "Bayesian-based automated model generation method for neural network modeling of microwave components," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 11, pp. 1179–1182, Nov. 2021.
- [160] B. Settles, "Active learning literature survey," Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep. 1648, 2010.
- [161] F. Olsson, "A literature survey of active machine learning in the context of natural language processing," Swedish Inst. Comput. Sci., Kista, Sweden, Tech. Rep. T2009:06, 2009.
- [162] C. Brezinski and M. R. Zaglia, *Extrapolation Methods: Theory and Practice*. New York, NY, USA: Elsevier, 1991.
- [163] L. C. Kiong, M. Rajeswari, and M. V. C. Rao, "Extrapolation detection and novelty-based node insertion for sequential growing multi-experts network," *Appl. Soft Comput.*, vol. 3, no. 2, pp. 159–175, Sep. 2003.
- [164] J. Xu, M. C. E. Yagoub, R. Ding, and Q.-J. Zhang, "Robust neural based microwave modelling and design using advanced model extrapolation," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2004, pp. 1549–1552.
- [165] L. Zhang and Q.-J. Zhang, "Simple and effective extrapolation technique for neural-based microwave modeling," *IEEE Microw. Wireless Compon. Lett.*, vol. 20, no. 6, pp. 301–303, Jun. 2010.
- [166] W. Na, W. Liu, L. Zhu, F. Feng, J. Ma, and Q.-J. Zhang, "Advanced extrapolation technique for neural-based microwave modeling and design," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 10, pp. 4397–4418, Oct. 2018.
- [167] J. W. Bandler *et al.*, "Space mapping: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.
- [168] H. M. Torun, A. C. Durgun, K. Aygun, and M. Swaminathan, "Causal and passive parameterization of S-parameters using neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 10, pp. 4290–4304, Oct. 2020.
- [169] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 1052–1061, Jul. 1999.
- [170] J. W. Bandler, M. A. Ismail, J. E. Rayas-Sánchez, and Q.-J. Zhang, "Neural inverse space mapping (NISM) optimization for EM-based microwave design," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 13, no. 2, pp. 136–147, Mar. 2003.
- [171] H. Yu, H. M. Torun, M. U. Rehman, and M. Swaminathan, "Design of SIW filters in D-band using invertible neural nets," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2020, pp. 72–75.
- [172] P. Jaraut *et al.*, "Augmented convolutional neural network for behavioral modeling and digital predistortion of concurrent multiband power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 9, pp. 4142–4156, Sep. 2021.
- [173] Z. Naghibi, S. A. Sadrossadat, and S. Safari, "Dynamic behavioral modeling of nonlinear circuits using a novel recurrent neural network technique," *Int. J. Circuit Theory Appl.*, vol. 47, no. 7, pp. 1071–1085, Jul. 2019.
- [174] J. Bandler, S. H. Chen, and S. Daijavad, "Microwave device modeling using efficient  $l_1$  optimization: A novel approach," *IEEE Trans. Microw. Theory Techn.*, vol. MTT-34, no. 12, pp. 1282–1293, Dec. 1986.
- [175] 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.
- [176] J. Xu, M. C. E. Yagoub, R. Ding, and Q. J. Zhang, "Exact adjoint sensitivity analysis for neural-based microwave modeling and design," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 1, pp. 226–237, Jan. 2003.
- [177] J. Wood, M. LeFevre, D. Runton, J.-C. Nanan, B. H. Noori, and P. H. Aaen, "Envelope-domain time series (ET) behavioral model of a Doherty RF power amplifier for system design," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 8, pp. 3163–3172, Aug. 2006.
- [178] J. Zhai, J. Zhou, L. Zhang, and W. Hong, "Behavioral modeling of power amplifiers with dynamic fuzzy neural networks," *IEEE Microw. Wireless Compon. Lett.*, vol. 20, no. 9, pp. 528–530, Sep. 2010.
- [179] F. Mkaem and S. Boumaiza, "Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 4, pp. 913–923, Apr. 2011.
- [180] S. Wang, M. Roger, J. Sarrazin, and C. Lelandais-Perrault, "Hyperparameter optimization of two-hidden-layer neural networks for power amplifiers behavioral modeling using genetic algorithms," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 12, pp. 802–805, Dec. 2019.
- [181] L. M. Aguilar-Lobo, J. R. Loo-Yau, S. Ortega-Cisneros, P. Moreno, and J. A. Reynoso-Hernandez, "Experimental study of the capabilities of the real-valued NARX neural network for behavioral modeling of multi-standard RF power amplifier," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2015, pp. 1–4.
- [182] S. Wang, M. Roger, J. Sarrazin, and C. Lelandais-Perrault, "Augmented iterative learning control for neural-network-based joint crest factor reduction and digital predistortion of power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 11, pp. 4835–4845, Nov. 2020.
- [183] P. Jaraut, M. Rawat, and F. M. Ghannouchi, "Composite neural network digital predistortion model for joint mitigation of crosstalk,  $I/Q$  imbalance, nonlinearity in MIMO transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 11, pp. 5011–5020, Nov. 2018.
- [184] A. Brihuega, L. Anttila, and M. Valkama, "Neural-network-based digital predistortion for active antenna arrays under load modulation," *IEEE Microw. Wireless Compon. Lett.*, vol. 30, no. 8, pp. 843–846, Aug. 2020.
- [185] J. Renteria-Cedano *et al.*, "A novel configurable FPGA architecture for hardware implementation of multilayer feedforward neural networks suitable for digital pre-distortion technique," in *Proc. 46th Eur. Microw. Conf. (EuMC)*, Oct. 2016, pp. 854–857.
- [186] E. Guillena, W. Li, G. Montoro, R. Quaglia, and P. L. Gilibert, "Reconfigurable DPD based on ANNs for wideband load modulated balanced amplifiers under dynamic operation from 1.8 to 2.4 GHz," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 1, pp. 453–465, Jan. 2022.
- [187] G. Avolio, A. Raffo, M. Marchetti, G. Bosi, V. Vadala, and G. Vannini, "GaN FET load-pull data in circuit simulators: A comparative study," in *Proc. 14th Eur. Microw. Integr. Circuits Conf. (EuMIC)*, Sep. 2019, pp. 80–83.

- [188] D. M. M. P. Schreurs, J. Verspecht, S. Vandenberghe, and E. Vandamme, "Straightforward and accurate nonlinear device model parameter-estimation method based on vectorial large-signal measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 10, pp. 2315–2319, Oct. 2002.
- [189] Z. Marinković *et al.*, "Neural approach for temperature-dependent modeling of GaN HEMTs," *Int. J. Numer. Model. Electron. Netw. Devices Fields*, vol. 28, pp. 359–370, Jun. 2014.
- [190] J. Xu, R. Jones, S. A. Harris, T. Nielsen, and D. E. Root, "Dynamic FET model—*DynaFET*—For GaN transistors from NVNA active source injection measurements," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2014, pp. 1–3.
- [191] A. Jarndal, "On neural networks based electrothermal modeling of GaN devices," *IEEE Access*, vol. 7, pp. 94205–94214, 2019.
- [192] X. Du *et al.*, "ANN-based large-signal model of AlGaIn/GaN HEMTs with accurate buffer-related trapping effects characterization," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 7, pp. 3090–3099, Jul. 2020.
- [193] O. Jardel *et al.*, "An electrothermal model for AlGaIn/GaN power HEMTs including trapping effects to improve large-signal simulation results on high VSWR," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 12, pp. 2660–2669, Dec. 2007.
- [194] S. Halder, J. McMacken, N. Craig, and J. Gering, "A hybrid modeling approach for drain-source capacitance of source field plated GaN FET devices for multi-bias application," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 1787–1790.
- [195] Z. Zhao, L. Zhang, F. Feng, W. Zhang, and Q.-J. Zhang, "Space mapping technique using decomposed mappings for GaN HEMT modeling," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 8, pp. 3318–3341, Aug. 2020.
- [196] W. Hu, H. Luo, X. Yan, and Y.-X. Guo, "An accurate neural network-based consistent gate charge model for GaN HEMTs by refining intrinsic capacitances," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 7, pp. 3208–3218, Jul. 2021.
- [197] 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.
- [198] Y. Cao, S. Reitzinger, and Q.-J. Zhang, "Simple and efficient high-dimensional parametric modeling for microwave cavity filters using modular neural network," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 5, pp. 258–260, May 2011.
- [199] S. Li, Y. Wang, M. Yu, and A. Panariello, "Efficient modeling of Ku-band high power dielectric resonator filter with applications of neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 8, pp. 3427–3435, Aug. 2019.
- [200] Y. Wang, M. Yu, H. Kabir, and Q. Zhang, "Effective design of cross-coupled filter using neural networks and coupling matrix," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2006, pp. 1431–1434.
- [201] S. Q. Li, Y. Wang, and M. Yu, "Effective modeling of filters in the design and optimization of multiplexers," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2015, pp. 1–3.
- [202] A. A. Ilumoka, "Efficient and accurate crosstalk prediction via neural net-based topological decomposition of 3-D interconnect," *IEEE Trans. Adv. Packag.*, vol. 24, no. 3, pp. 268–276, Aug. 2001.
- [203] K. Hsu, W. Guo, G. Shiue, C. Lin, T. Huang, and R. Wu, "Design of reflectionless vias using neural network-based approach," *IEEE Trans. Adv. Packag.*, vol. 31, no. 1, pp. 211–218, Feb. 2008.
- [204] T. Lu, J. Sun, K. Wu, and Z. Yang, "High-speed channel modeling with machine learning methods for signal integrity analysis," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 6, pp. 1957–1964, Dec. 2018.
- [205] C.-C. Chiu, K.-Y. Yang, Y.-H. Lin, W.-S. Wang, T.-Y. Wu, and R.-B. Wu, "A novel dual-sided fly-by topology for 1–8 DDR with optimized signal integrity by EBG design," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 8, no. 10, pp. 1823–1829, Oct. 2018.
- [206] M. Swaminathan, H. M. Torun, H. Yu, J. A. Hejase, and W. D. Becker, "Demystifying machine learning for signal and power integrity problems in packaging," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 10, no. 8, pp. 1276–1295, Aug. 2020.
- [207] Y. Tian, F. He, Q.-J. Zhang, C. M. Tan, and J. Ma, "Rapid ULSI interconnect reliability analysis using neural networks," *IEEE Trans. Device Mater. Rel.*, vol. 14, no. 1, pp. 400–407, Mar. 2014.
- [208] H. Yin, Z. Jiang, X.-W. Zhu, and C. Yu, "Over-the-air behavioral modeling of millimeter wave beamforming transmitters with concurrent dynamic configurations utilizing heterogeneous neural network," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2020, pp. 397–400.
- [209] W. Zhang, A. Hasan, F. M. Ghannouchi, M. Helaoui, Y. Wu, and Y. Liu, "Novel calibration algorithm of multipoint wideband receivers based on real-valued time-delay neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 11, pp. 3540–3548, Nov. 2016.
- [210] P. H. Aaen *et al.*, "Multiphysics modeling of RF and microwave high-power transistors," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 12, pp. 4013–4023, Dec. 2012.
- [211] M. Gall, M. Gardill, J. Fuchs, and T. Horn, "Learning representations for neural networks applied to spectrum-based direction-of-arrival estimation for automotive radar," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2020, pp. 1031–1034.
- [212] S. Jeong, T.-H. Lin, and M. M. Tentzeris, "A real-time range-adaptive impedance matching utilizing a machine learning strategy based on neural networks for wireless power transfer systems," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 12, pp. 5340–5347, Dec. 2019.
- [213] S. Wei, Q. Qu, X. Zeng, J. Liang, J. Shi, and X. Zhang, "Self-attention Bi-LSTM networks for radar signal modulation recognition," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 11, pp. 5160–5172, Nov. 2021.
- [214] E. Klinefelter and J. A. Nanzer, "Automotive velocity sensing using millimeter-wave interferometric radar," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 1, pp. 1096–1104, Jan. 2021.



**Feng Feng** (Member, IEEE) received the B.Eng. degree from Tianjin University, Tianjin, China, in 2012, and the Ph.D. degrees from the School of Microelectronics, Tianjin University, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2017.

From 2017 to 2020, he was a Post-Doctoral Fellow with the Department of Electronics, Carleton University. He is currently an Associate Professor with the School of Microelectronics, Tianjin University. His research interests include electromagnetic parametric modeling and design optimization algorithms, deep neural network modeling method, space mapping algorithm and surrogate model optimization, finite-element method in electromagnetic simulation and optimization, and multi-physics modeling and optimization.



**Weicong Na** (Member, IEEE) received the B.Eng. degree from Tianjin University, Tianjin, China, in 2012, and the Ph.D. degrees from the School of Microelectronics, Tianjin University, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2018.

She is currently with the Faculty of Information Technology, Beijing University of Technology, Beijing, China. Her research interests include microwave circuit modeling and design, automated neural network model generation algorithm, electromagnetic (EM) field knowledge-based modeling and optimization, and deep neural network modeling for microwave applications.



**Jing Jin** (Member, IEEE) received the B.Eng. degree from Wuhan University, Wuhan, China, in 2014, and the Ph.D. degrees from the School of Microelectronics, Tianjin University, Tianjin, China, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2021.

She is currently an Associate Professor with the College of Physical Science and Technology, Central China Normal University, Wuhan. Her research interests include microwave circuit modeling and design, deep neural network technique for microwave modeling, surrogate-based electromagnetic (EM) optimization, and EM topology optimization.



**Jianan Zhang** (Member, IEEE) received the B.Eng. degree from Tianjin University, Tianjin, China, in 2013, and the Ph.D. degrees from the School of Microelectronics, Tianjin University, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2020.

From 2020 to 2022, he was a Post-Doctoral Research Associate with the Department of Electronics, Carleton University. He is currently an Associate Professor with the School of Information Science and Engineering, Southeast University, Nanjing, China. His research interests include neural network-based electromagnetic (EM) parametric modeling and optimization, surrogate modeling and surrogate-assisted optimization, finite-element analysis in EM, and quantum computing with applications to EM problems.



**Wei Zhang** (Member, IEEE) received the B.Eng. degree from Shandong University, Shandong, China, in 2013, and the Ph.D. degrees from the School of Microelectronics, Tianjin University, Tianjin, China, and the Department of Electronics, Carleton University, Ottawa, ON, Canada, in 2020.

From 2020 to 2021, he was a Post-Doctoral Fellow with the Department of Electronics, Carleton University. He is currently an Associate Researcher with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China.

His research interests include neural-network-based methods for microwave device modeling, electromagnetic centric multiphysics modeling and optimization, space mapping algorithm and surrogate model optimization, and computational electromagnetics.



**Qi-Jun Zhang** (Fellow, IEEE) received the B.Eng. degree from the Nanjing University of Science and Technology, Nanjing, China, in 1982, and the Ph.D. degree in electrical engineering from Mc-Master University, Hamilton, ON, Canada, in 1987.

From 1982 to 1983, he was with the System Engineering Institute, Tianjin University, Tianjin, China. From 1988 to 1990, he was with Optimization Systems Associates (OSA) Inc., Dundas, ON, Canada, where he developed advanced microwave optimization software. In 1990, he joined the Department of Electronics, Carleton University, Ottawa, ON, Canada, where he is currently a Chancellor's Professor. He has authored or coauthored over 300 publications. He authored *Neural Networks for RF and Microwave Design* (Artech House, 2000), coedited *Modeling and Simulation of High-Speed VLSI Interconnects* (Kluwer, 1994) and *Simulation-Driven Design Optimization and Modeling for Microwave Engineering* (Imperial College Press, 2013), and contributed to the *Encyclopedia of RF and Microwave Engineering* (Wiley, 2005), *Fundamentals of Nonlinear Behavioral Modeling for RF and Microwave Design* (Artech House, 2005), and *Analog Methods for Computer-Aided Analysis and Diagnosis* (Marcel Dekker, 1988). He was a Guest Coeditor for the Special Issue on High-Speed VLSI Interconnects for the *International Journal of Analog Integrated Circuits and Signal Processing* (Kluwer, 1994), and twice was a Guest Editor for the Special Issue on Applications of ANN to RF and Microwave Design for the *International Journal of RF and Microwave Computer-Aided Engineering* (Wiley, 1999 and 2002). His research interests include microwave design automation, especially on neural-network and optimization methods for high-speed/high-frequency circuit design.

Dr. Zhang is a Fellow of the Canadian Academy of Engineering. He is the Chair of the Technical Committee on Design Automation (MTT-2) of the IEEE Microwave Theory and Techniques Society (IEEE MTT-S). He is an Associate Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES.