

# A Solution to the Dilemma for FSS Inverse Design Using Generative Models

Zheming Gu<sup>✉</sup>, Da Li<sup>✉</sup>, Member, IEEE, Yunlong Wu<sup>✉</sup>, Yudi Fan<sup>✉</sup>, Chengting Yu, Hongsheng Chen<sup>✉</sup>, Fellow, IEEE, and Er-Ping Li<sup>✉</sup>, Fellow, IEEE

**Abstract**—Recently, artificial neural networks (ANNs) show a great potential in frequency-selective surface (FSS) inverse design. However, it is inevitable to encounter the problem of nonunique mapping between inputs and outputs, which cannot be easily solved by the traditional ANNs framework. We analyze this existing dilemma from the perspective of information loss caused by data dimensionality reduction and propose deploying generative models as a solution for the first time. Specifically, two approaches with a novel model based on conditional generative adversarial network (cGAN) are presented to achieve inverse design from the given indexes to FSS physical dimensions. By applying the proposed method, we can immediately obtain the FSS design that meets the industrial demands without complex neural network processing or repeated iterations. Moreover, the proposed method is validated in closed-loop simulations and corresponding experiments, which also paves the way for designing complex FSS structures with the desired electromagnetic responses using deep neural networks.

**Index Terms**—Frequency-selective surface (FSS), generative adversarial network (GAN), generative models, inverse design, nonuniqueness.

## I. INTRODUCTION

FORWARD design methods, such as full-wave simulation and equivalent circuit analysis, have been well studied for decades to analyze the electromagnetic performance of

Manuscript received 28 September 2022; revised 20 March 2023; accepted 1 April 2023. Date of publication 14 April 2023; date of current version 2 June 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62201499, Grant 62071424, and Grant 62027805; and in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LQ23F010019 and Grant LD21F010002. (Zheming Gu and Yunlong Wu contributed equally to this work.) (Corresponding author: Da Li.)

Zheming Gu, Yunlong Wu, Yudi Fan, Chengting Yu, Hongsheng Chen, and Er-Ping Li are with the Key Laboratory of Advanced Micro/Nano Electronic Devices and Smart Systems and Applications, College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and also with the Zhejiang University–University of Illinois at Urbana–Champaign Institute, Zhejiang University, Haining 314400, China (e-mail: guzheming@zju.edu.cn; wuyunlong@zju.edu.cn; fanyudi@zju.edu.cn; chengting.21@intl.zju.edu.cn; hansomchen@zju.edu.cn; liep@zju.edu.cn).

Da Li is with the Key Laboratory of Advanced Micro/Nano Electronic Devices and Smart Systems and Applications, College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, also with the Zhejiang University–University of Illinois at Urbana–Champaign Institute, Zhejiang University, Haining 314400, China, and also with the Zhejiang–Singapore Innovation and AI Joint Research Laboratory, Hangzhou 310027, China (e-mail: li-da@zju.edu.cn).

Data is available on-line at <https://github.com/ushione/Generative-Models-for-FSS-Inverse-Design>.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAP.2023.3266053>.

Digital Object Identifier 10.1109/TAP.2023.3266053

frequency-selective surface (FSS) structures [1], [2], [3], [4], [5], [6]. In reality, the most critical problem is how to quickly determine the physical sizes to meet the expected design requirements. Although the iteration and trial-and-error of this inverse design process can be accelerated by evolutionary algorithms and others, the large consumption of time and computing resources is still a severe problem.

In recent years, artificial neural networks (ANNs) have been regarded as powerful tools in the electromagnetic field because of their nonlinear characterization ability [7], [8], [9], [10], [11]. Forward optimization methods have been further developed by building well-trained forward neural networks as efficient equivalents for simulation [12], [13], [14], [15], [16]. Direct inverse modeling is further proposed as a more efficient design method, which can provide the physical parameters in only one assessment without iterations [17], [18], [19], [20], [21], [22]. However, unlike forward optimization methods that infer electromagnetic performance from physical parameters and solve a one-to-one mapping, there are often multiple solutions in the inverse design to meet a given design requirements. Therefore, a challenge of inverse design is that the nonunique mapping from input to output makes it difficult to train and apply neural networks [20].

To alleviate this problem, Kabir et al. [19] divided the dataset into groups and trained multiple networks to combine them into an inverse model. However, this approach is relatively ineffective, especially when faced with a large dataset, since eliminating the apparent conflicting instances does not fundamentally address the issue of nonunique mapping [20], [21]. Zhang et al. [21] further proposed to extend the neural network output dimension to allow multivalued mapping. However, this method is sensitive and fragile, requiring careful adjustment of extended dimension size. Zhu et al. [22] adopted a more specific requirement description as the network input to alleviate the nonunique problem. However, such an inverse network becomes challenging to apply to practical engineering because of the gap between its input and the realistic design requirements. Although these studies attempt to improve discriminative neural networks to learn nonunique mapping tasks from the aspects of dataset preprocessing, network output dimension, network input dimension, and so on, the inherent flaws of such neural networks make the inverse design still not easy.

The main contributions of this article are summarized as follows.

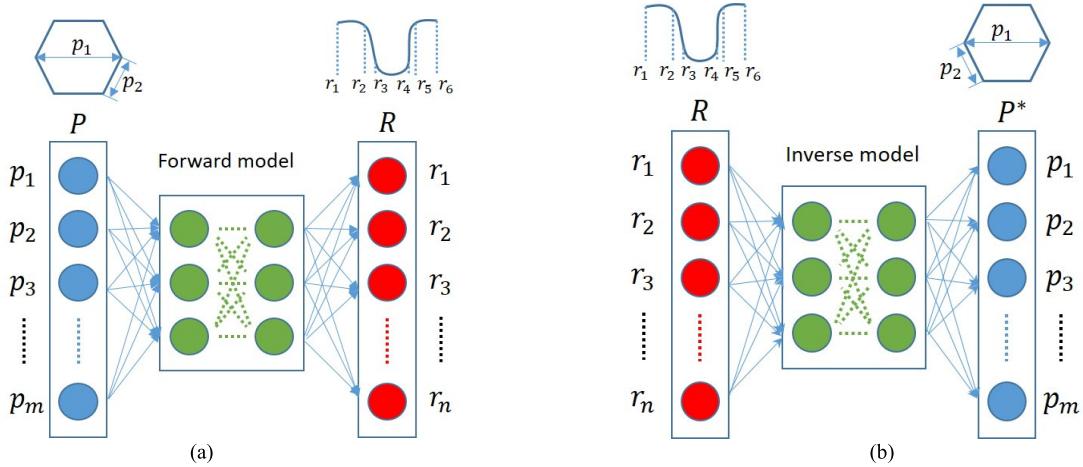


Fig. 1. (a) Forward model takes physical parameters  $P$  as inputs and corresponding responses  $R$  as outputs. (b) Inverse model takes required restraints  $R$  as inputs and physical parameters  $P^*$  as outputs.

1) For the first time, we formulated and discussed the problem of nonunique mapping from the new perspective of information loss caused by data dimensionality reduction and pointed out the limitations of discriminative models in the FSS inverse design.

2) Instead of improving the discriminative models, we propose to apply generative models as a solution to this dilemma, which introduces latent variables to model the missing information. We also give an instance of using the generative model to FSS inverse design, where the physical parameters that make the FSS to meet the desired electromagnetic performance can be solved immediately.

3) Furthermore, we propose and open source two specific implementations based on conditional generative adversarial networks (cGANs) for FSS inverse design. One is based on the well-trained inverse network, and the other is an end-to-end paradigm. The former establishes connections with existing inverse design work, and the latter paves a new path for a series of inverse design problems. The effectiveness of both has also been verified in simulation and testing.

The proposed method is more efficient and advantageous than the forward optimization method because it can immediately provide a solution that meets the design requirements without iterations. This advantage is further amplified when multiple optimization requirements are encountered in real engineering. Compared with the direct inverse modeling methods, a new and flexible network architecture is proposed to address the challenge of nonunique mappings. In addition, the proposed method accepts more realistic design requirements and delivers multiple alternatives simultaneously, which is essential for inverse design engineering.

The structure of this article is organized as follows. Section II discusses the design dilemma of inverse design, and the core idea of our proposed method is discussed in detail. In Section III, the inverse design of FSS is completed by deploying a cGAN to exploit the well-trained inverse network. Another end-to-end solution is proposed and validated in Section IV. In Section V, we compare the proposed method with other methods.

Experimental verification is carried out in Section VI. Conclusions are finally drawn in Section VII.

## II. BACKGROUND

In this section, we first formulate and discuss the inverse design dilemma caused by the nonuniqueness problem. Then, we interpret this dilemma from the perspective of information loss caused by data compression and propose deploying generative adversarial networks (GANs) (a generative model) as a favorable solution. The preliminaries of GAN are also introduced in this section.

### A. Dilemma for Microwave Inverse Design

Neural networks have been shown to predict the monitored electromagnetic responses  $R$  from physical parameters  $P$ . Optimization methods in microwave inverse design have also been further developed, where well-trained neural networks play an efficient equivalent to simulation [12], [13], [14], [15], [16]. However, iterative optimization of physical parameters  $P$  to obtain the desired electromagnetic response  $R$  is still indispensable during this process. Therefore, as shown in Fig. 1, the direct inverse modeling design is further proposed, in which neural networks adopt swapped input–output pairs ( $R \rightarrow P$ ) compared to the forward model.

Some cognitive-based descriptions (e.g., start and stop frequencies of passband or stopband) are realistic metrics of the monitored electromagnetic response, denoted as  $R_{low}$ , since they are highly abstract. However, training such a neural network for inverse design may present some challenges due to the nonuniqueness of the data. One fact is that different physical parameters may produce similar or even identical monitored electromagnetic responses, which leads to the presence of input–output pairs such as  $(R'_{low} \rightarrow P^A)$  and  $(R'_{low} \rightarrow P^B)$  in the dataset. This problem becomes more severe when such a lower dimensional description is used as the network input, leading to a decrease in network accuracy or even failure to converge [19], [20], [21], [22].

Therefore, some high-dimensional descriptions (e.g., S-parameter curves) are used more often as the

network input, denoted as  $R_{high}$ , since they contain more detailed descriptions of the electromagnetic response. The conflicting data in the dataset are significantly reduced when using  $R_{high}$ , which can mitigate the nonunique problem well. However, a new dilemma arises when such inverse networks are used for practical inverse designs. Industrial demands (e.g., passband and stopband) are ambiguous compared to S-parameter curves, so it is not easy to translate them into inputs to the network. An obvious example is that the S-parameter curve that satisfies the demands may not be achievable for the current FSS structure, so the trained neural network naturally cannot give feasible physical parameters. This is the key problem in the microwave inverse design discussed in this article.

### B. Novel Solution to the Dilemma Based on Generative Models

We propose that the nonunique mapping issue is actually caused by the partial absence of features. Considering that the given physical parameter  $P$  describes a unique structure and further determines the monitored electromagnetic response  $R$ ,  $\{P\} \rightarrow \{R\}$  is deemed to satisfy the mapping relationship. However,  $\{R\} \rightarrow \{P\}$  may not, especially when  $R_{low}$  is used. We further argue that while  $P$  determines the electromagnetic response  $R$ , some hidden variables  $H$  are also determined simultaneously (e.g., the part of information lost from  $R_{high}$  to  $R_{low}$ ).

The deep neural networks deployed in the above-related work [20], [21], [22] are also known as discriminative models, which map the given unobserved variable  $x$  to a label  $y$  dependent on the training samples data or model a conditional probability distribution  $P(y|x)$ . Therefore, applying the discriminative models in the inverse design may fail when the latent variable  $H$  is present.

In this article, we propose a solution to this dilemma by using a generative model, which studies the joint probability  $P(x, y)$ . One advantage of using a generative model is that an additional variable  $z$  can be introduced to model the latent variable  $H$  to realize the mapping relationship of  $\{R, H\} \rightarrow \{P\}$ , which solves the problem of nonunique data.

### C. Preliminaries of Generative Adversarial Networks

As a generative model, GAN has been widely used in data enhancement, image generation, image translation, and other fields due to its excellent performance in recent years [23]. There are generally two networks in a GAN: a *generator*  $G$  that relies on random noise  $z$  to generate data  $\tilde{x}$  similar to the target data  $x$  and a *discriminator*  $D$  that distinguishes the fake data  $\tilde{x}$  from real data  $x$  as much as possible. In other words,  $G$  and  $D$  play the following two-player minimax game with value function  $V(D, G)$  [23]:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data(x)}} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

Although vanilla GANs work well for generative tasks, further employment is limited due to the uncontrollable output.

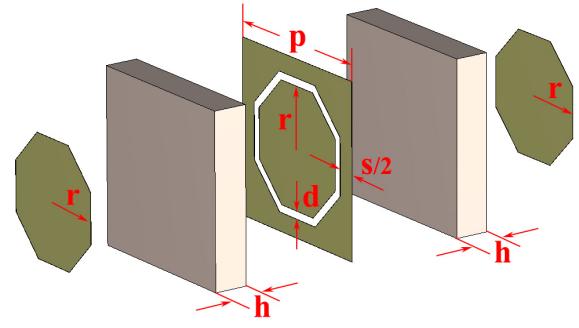


Fig. 2. Schematic of the three-layer FSS model.

TABLE I  
SPECIFIC FSS STRUCTURE PARAMETERS

Params	Range (mm)	Step (mm)
$h$	2.00~5.00	0.25
$d$	0.15~0.65	0.10
$s$	2.90~3.30	0.10
$r$	2.50~9.10	0.30

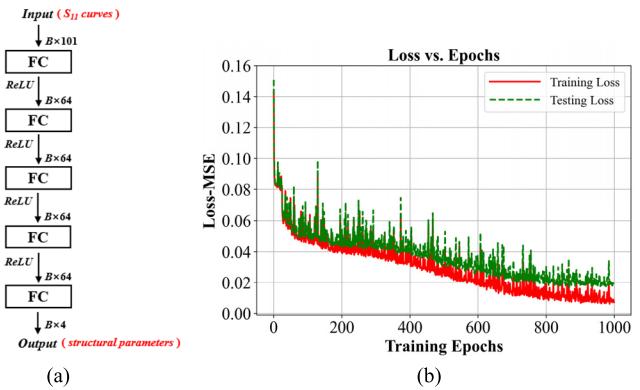


Fig. 3. (a) Architecture diagram of the inverse network. After the 101 sampling points of the  $S_{11}$  curve are fed, the inverse network outputs the corresponding FSS structure parameters. Note that ReLU is used as the activation layer after the fully connected layer, while the data flow with the batch size  $B$  is also illustrated. (b) Loss curves of the inverse network during training.

Therefore, cGAN is proposed to constrain the output by further introducing a new semantic  $c$ . The objective function of a two-player minimax game is described as [24]

$$\begin{aligned} \min_G \max_D V(D, G) = & E_{x \sim p_{data(x)}} [\log D(x|c)] \\ & + E_{z \sim p_z(z)} [\log(1 - D(G(z|c)))] \end{aligned} \quad (2)$$

### III. INVERSE DESIGN BY JOINT INVERSE NETWORK AND GAN

In this section, an instance of FSS inverse design is introduced, which achieves the desired electromagnetic response by controlling several physical parameters of the FSS unit. An inverse network is first built to output the physical parameters, which takes  $S_{11}$  as input to avoid the nonunique problem. Then, a well-designed cGAN is utilized to achieve translation

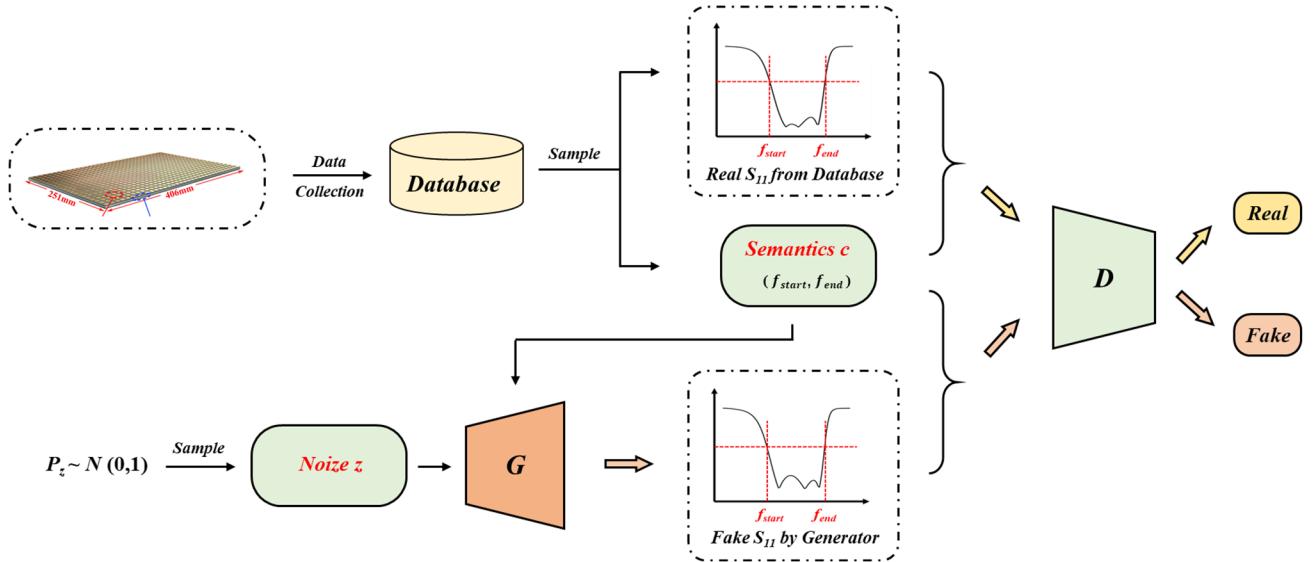


Fig. 4. Architecture of cGAN model.

from semantic  $c$  to  $S_{11}$ , which solves the dilemma of applying inverse networks. Finally, the effectiveness of the proposed method in the inverse design problem is proved by simulation.

#### A. Design Example of FSS Inverse Design

As shown in Fig. 2, the FSS unit is a three-layer structure. Its physical parameters can be expressed as  $h$ ,  $d$ ,  $s$ , and  $r$ . The period of the unit is set as  $p = 2r + s + 0.65$ , and the relative dielectric constant of substrates is 2.2. The full-wave simulation was used to scan physical parameters according to Table I, while the reflection coefficient ( $S_{11}$ ) and the transmission coefficient ( $S_{21}$ ) were also recorded. Finally, 7144 pieces of valid data were generated by the CST Studio Suite Frequency Domain Solver in this study. Furthermore,  $-10$  dB is used as a threshold to determine the passband by observing  $S_{11}$ . Note that the broadest passband is primarily concerned in this work.

As described in Section II, to realize the inverse design of the FSS structure, the inverse network with the architecture shown in Fig. 3 is trained using 5000 pieces of data, and its performance is further verified on the remaining 2144 pieces of data. Note that the  $S_{11}$  curve containing 101 points of equally spaced frequency samples is fed to the inverse network as a high-dimensional electromagnetic response description  $R_{high}$ , which avoids the problem of nonunique data.

#### B. Implementation of the cGAN Model

It should be noted that the desired passband demands must be translated into a reliable  $S_{11}$  curve to use the above well-trained inverse network. Here, a cGAN model with the architecture shown in Fig. 4 is deployed to accomplish this goal. Fig. 5 further shows the details of the generator and discriminator deployment.

The cGAN model accepts an input consisting of two parts: a Gaussian noise  $z$  sampled from a standard normal distribution and an embedding vector expressing the semantics of the passband requirements  $c$ . The  $S_{11}$  curve satisfies the demands

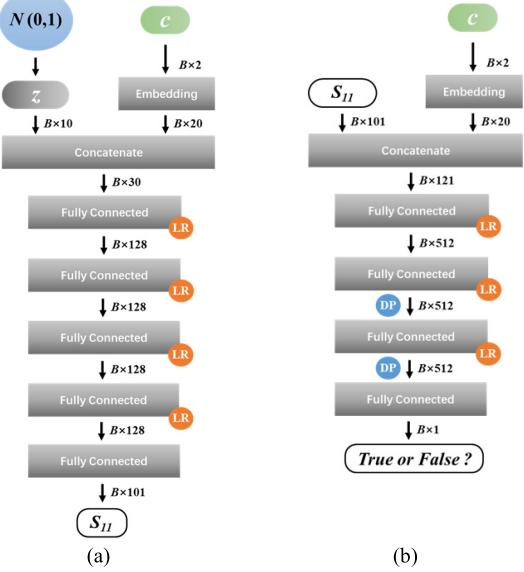


Fig. 5. Architecture diagram of (a) generator  $G$  and (b) discriminator  $D$ . Note that the LR annotation indicates that LeakyReLU is used as the activation function after this fully connected layer. The DP annotation indicates connecting dropout layers to suppress overfitting. The data stream with batch training size  $B$  is also given.

by specifying the semantic  $c$ , while the noise  $z$  models the latent variables and guarantees the diversity of the curve. The critical steps and procedures in the cGAN model development are clarified in Algorithm 1. The following two criteria are used to judge whether the GAN converges: 1) the probability that the discriminator distinguishes the generator is about 50%, which indicates that the distribution of the generated and actual data is similar and 2) the output of the generator can withstand in Section III-C.

To finally obtain the physical parameters corresponding to the desired passband performance, it is also necessary to use the curve generated by the GAN as the input of the inverse network. However, since it is difficult for the fully connected layer to grasp the local information of the curve,

$$\frac{1}{N} \times \left\{ W_{Re} \times \begin{pmatrix} m \times m \\ \begin{matrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{matrix} \\ \vdots \\ \begin{matrix} 0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{matrix} \end{pmatrix} O \right\}_{N \times (2M-1)} \times \begin{pmatrix} x_{FC}^0 \\ x_{FC}^1 \\ \vdots \\ x_{FC}^m \\ x_{FC}^{m+1} \\ \vdots \\ x_{FC}^{2M-2} \end{pmatrix}_{(2M-1) \times 1} - W_{Im} \times \left\{ O \times \begin{pmatrix} 0 & \cdots & 0 \\ \vdots \\ \begin{matrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{matrix} \\ \vdots \\ \begin{matrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{matrix} \end{pmatrix}_{(m-1) \times (m-1)} \right\}_{N \times (2M-1)} \times \begin{pmatrix} x_{FC}^0 \\ x_{FC}^1 \\ \vdots \\ x_{FC}^m \\ x_{FC}^{m+1} \\ \vdots \\ x_{FC}^{2M-2} \end{pmatrix}_{(2M-1) \times 1} = S_{11}$$

Fig. 6. Specific details of IDFT by building sparse connections in the Fourier layer.

the curve generated by the GAN is not smooth enough, which may lead to the degradation of its quality as the input of the inverse network [25]. To solve this problem, we propose to embed the Fourier transform into the neural network so that the smoothness of the curve can be adjusted by controlling the Fourier series to remove high-frequency components.

To be precise, a Fourier layer is added in the last layer of the generator, which accepts the output from the previous fully connected layer and performs a similar function to the inverse discrete Fourier transform (IDFT). Equation (5) gives the definition of IDFT

$$Y[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{j \cdot \frac{2\pi}{N} \cdot k \cdot n}. \quad (3)$$

Its matrix form can be expressed as

$$\frac{1}{N} \cdot \begin{bmatrix} w^{0 \cdot 0} & w^{0 \cdot 1} & \cdots & w^{0 \cdot (N-1)} \\ w^{1 \cdot 0} & w^{1 \cdot 1} & \cdots & w^{1 \cdot (N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ w^{(N-1) \cdot 0} & w^{(N-1) \cdot 1} & \cdots & w^{(N-1) \cdot (N-1)} \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_{N-1} \end{bmatrix} \quad (4)$$

$$w = e^{j \cdot \frac{2\pi}{N}} = \cos\left(\frac{2\pi}{N}\right) + \sin\left(\frac{2\pi}{N}\right) \cdot j. \quad (5)$$

$Y[n]$  is used to describe the real curve and represent an  $S_{11}$  curve composed of  $N$  points, where  $N$  is set to 101 in this work; thus,  $X[n]$  has a conjugate property. Considering the difficulty of deploying complex tensors in neural networks, it is necessary to decompose them into two real tensors, the real and the imaginary.

In addition, to control the number of first  $M$  terms of the Fourier series to filter high-frequency components, the last fully connected layer is reconstructed into  $2M - 1$  neurons, which in turn represent the fundamental frequency, the real part of the first  $M - 1$  Fourier coefficients and the imaginary part of the first  $M - 1$  Fourier coefficients. Therefore, the final output vector  $O_{FL}$  ( $S_{11}$ ) is calculated as follows:

$$\frac{1}{N} \cdot (W_{Re} \cdot X_{Re}[n] - W_{Im} \cdot X_{Im}[n]) = Y_{Re}[n] = O_{FL} \quad (6)$$

where subscripts  $Re$  and  $Im$  represent the real and imaginary parts, respectively. Furthermore, the mapping  $X_{FC}$  to  $X[n]$  is achieved by constructing a sparse connection relationship, and the implementation of IDFT in the neural network is shown in Fig. 6.

Note that as a neural network layer, the Fourier layer is parameter-free and sparsely connected. To further observe the progress of backpropagation, the *Jacobian matrix* between the input and the output of the Fourier layer is calculated as follows:

$$\frac{\partial(y_0, \dots, y_{N-1})}{\partial(x_0, \dots, x_{2M-2})} = \frac{1}{N} \cdot \begin{bmatrix} \cdots & & \cdots \\ \cdots & \frac{\partial y_i}{\partial x_j} & \cdots \\ \cdots & & \cdots \end{bmatrix} \quad (7)$$

where  $\frac{\partial y_i}{\partial x_j}$  is defined as

$$\frac{\partial y_i}{\partial x_j} = \begin{cases} 1 & j = 0; \\ \cos\left(\frac{2\pi \cdot i \cdot j}{N}\right) + \cos\left(\frac{2\pi \cdot i \cdot (N-j)}{N}\right) & 1 \leq j \leq M-1 \\ \sin\left(\frac{2\pi \cdot i \cdot (N-j+M-1)}{N}\right) & M \leq j. \\ -\sin\left(\frac{2\pi \cdot i \cdot (j-M+1)}{N}\right) \end{cases} \quad (8)$$

It is observed that there are exact upper and lower bounds on the gradient of the Fourier layer and help generate the curve by adjusting the different frequency components during the backpropagation process. With the introduction of a Fourier layer, the task of drawing the curve turns to determine the discrete Fourier series, which reduces the correlation between the outputs to help draw smoother curves. In addition, the number of terms of the discrete Fourier series is also a powerful means to limit the smoothness of the curve.

### C. Simulation Verification

The validity of the algorithm needs to be verified through the following three aspects.

1) *Precise Controllability*: The curve generated by GAN can be effectively controlled by different semantics.

2) *Generating Diversity*: The curves generated by GAN are diverse.

3) *Result Feasibility*: The GAN-generated curve, which characterizes the electromagnetic response, can be realized by the actual structure with appropriate parameters.

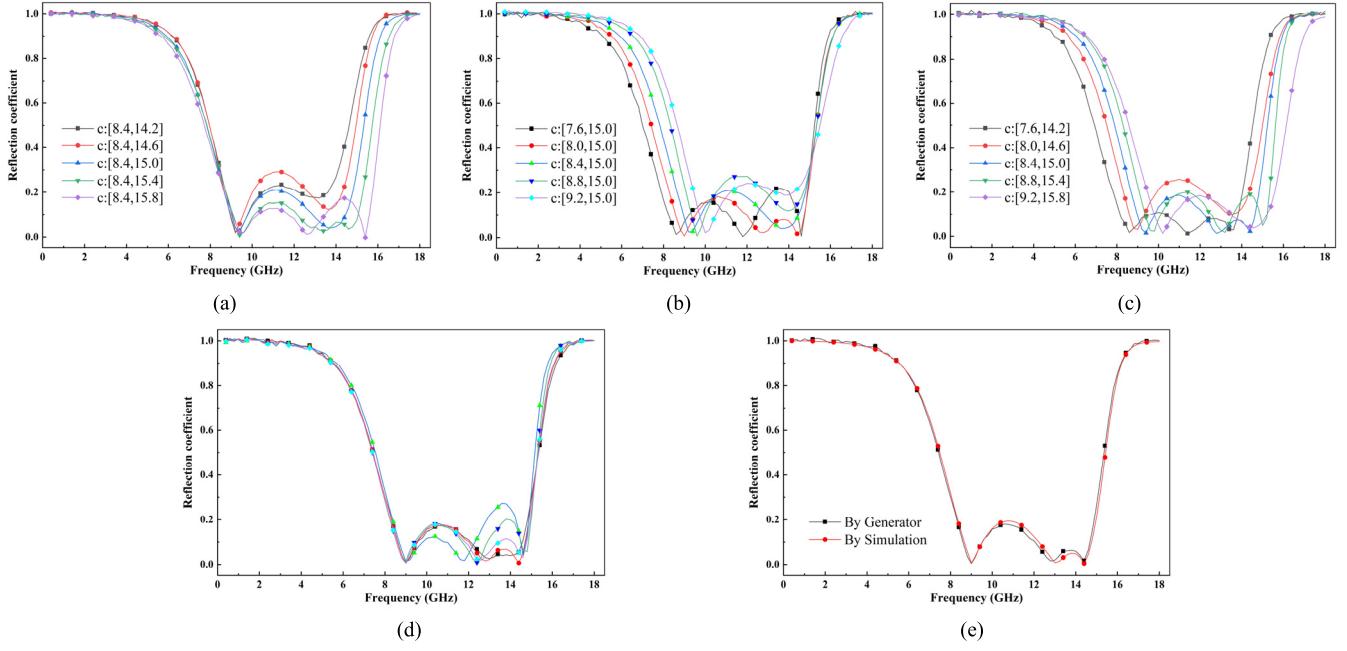


Fig. 7. Verify that the well-trained GAN learns the low-dimensional manifold of  $S_{11}$  in simulation. Note that in semantic representation verification [(a)–(c)], the GAN receives a fixed noise vector  $z$  and a gradually changing passband semantics  $c$ . (a) Fix the starting frequency of the passband and gradually adjust the cutoff frequency. (b) Fix the cutoff frequency of the passband and gradually change the start frequency. (c) Shift passband but fix its width. Validation (d) shows the diversity of solutions given by GANs, where GANs accept a fixed desired passband [8.0, 15.0] and different random noises  $z$ . Verification (e) shows that the  $S_{11}$  generated by GAN can be realized by the actual FSS structure. One  $S_{11}$  curve in (d) is selected as the input of the well-trained inverse network, and the simulation result of the FSS structure corresponding to the output structural parameters is highly consistent with the curve generated by GAN.

To demonstrate whether the passband semantics  $c$  are accurately represented in the generated  $S_{11}$ , we incrementally adjust the passband semantics to observe the generated  $S_{11}$  by fixing the noise  $z$ . Specifically, Fig. 7(a) shows the corresponding curve generated by GAN receiving the gradient of passband cutoff frequency semantics when the first transmission pole is fixed. Then, in Fig. 7(b), the starting frequency of the passband is gradually changed, while the cutoff frequency is kept fixed. Finally, in Fig. 7(c), the passband shifts gradually but keeps the bandwidth fixed. Therefore, all the above results strongly indicate that the trained GAN learns the low-dimensional manifold of  $S_{11}$  well and realizes the conditional generation of  $S_{11}$  through passband semantics.

Fig. 7(d) further shows the diversity of GAN generation curves. Specifically, when the passband semantics  $c$  is fixed, the GAN receives different noises  $z$  and output curves of different styles satisfying the semantics. This suggests that generative models GANs effectively model hidden variables by introducing noise  $z$  and solve the one-to-many mapping task that discriminative models cannot achieve.

The last and most crucial point is that it needs to be proved that the  $S_{11}$  curve generated by GAN can be effectively implemented by the structure while satisfying the passband semantics. Thus, first, we take the GAN-generated curve as input into the inverse network to get a set of physical parameters. Then, we use the structure parameter to model the corresponding FSS structure and carry out full-wave simulation verification. Fig. 7(e) shows the  $S_{11}$  generated by the GAN and the simulated one, where the two curves are highly consistent.

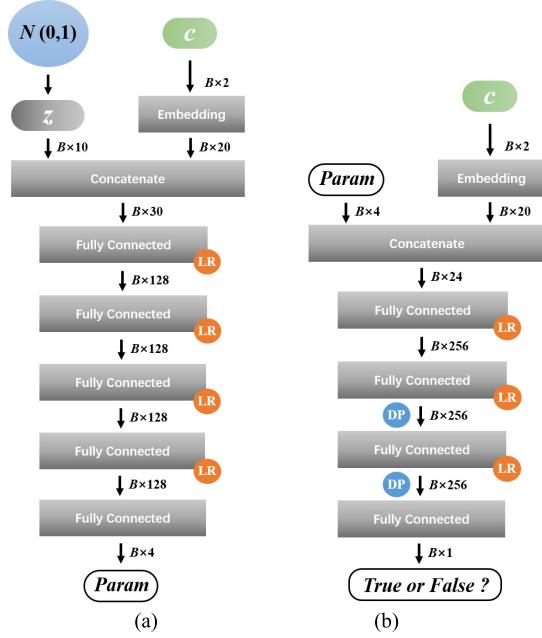


Fig. 8. Architecture diagram of (a) generator  $G$  in the end-to-end GAN and (b) discriminator  $D$  in the end-to-end GAN.

So far, the effectiveness of our proposed method has completed the closed-loop verification. Specifically, starting from the simulated dataset, we train an inverse model and propose to deploy a GAN to facilitate better use of the inverse model. Ultimately, the goal from the desired passband requirement to the determination of feasible physical parameters was achieved.

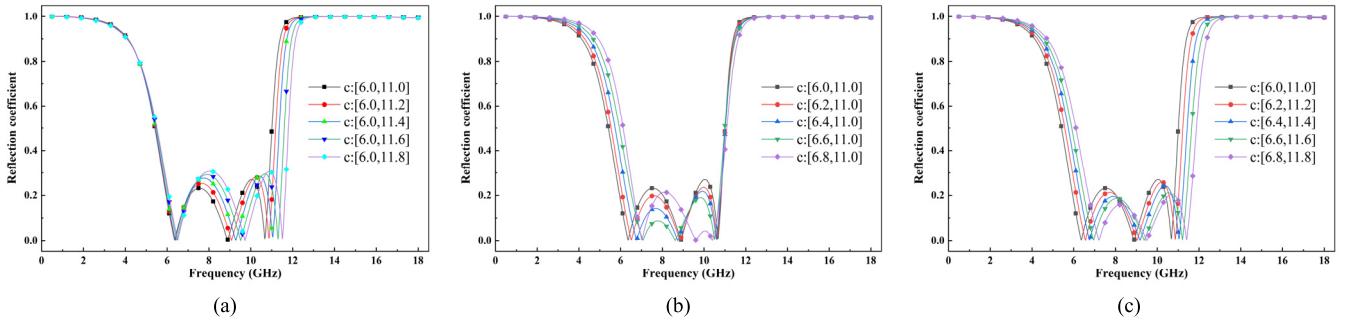


Fig. 9. End-to-end GAN learns the low-dimensional manifold of  $S_{11}$ . Note that in this comparative experiment, the GAN receives a fixed noise vector  $z$  and a gradually changing passband semantics  $c$ . (a) Fix the starting frequency and gradually adjust the cutoff frequency. (b) Fix the cutoff frequency and gradually adjust the starting frequency. (c) Shift passband while fixing its width.

#### IV. INVERSE DESIGN BY AN END-TO-END GAN

##### A. End-to-End GAN

In the previous part, GAN and an inverse network are used to accomplish the inverse design. First, GAN converts the demands into a feasible  $S_{11}$  curve, after which the inverse network takes this as input to obtain the final physical parameters. Since  $S_{11}$  curves are high-dimensional descriptions of electromagnetic responses  $R_{high}$ , mapping curves to physical parameters is considered achievable.

However, in the FSS design process, a set of physical parameters maps and determines its electromagnetic response. Also, this mapping relationship is completed by the trained inverse network in Section III. Considering that not all curves can be feasible  $S_{11}$  for the actual structure, the support set of the learning objective of the generator (i.e., the  $S_{11}$  curve) is actually a low-dimensional manifold in a high-dimensional space. We further argue that the mapping task of the inverse network can also be delegated to the GAN simultaneously, thereby building an end-to-end GAN model that directly models the joint probability of demands and physical parameters.

The training of the end-to-end GAN is similar to the previous section but uses the passband semantic-physical parameters as the dataset. The architecture of  $G$  and  $D$  in the end-to-end GAN is given in Fig. 8. The trained generator directly outputs the feasible solutions for the physical parameters that satisfy the desired passband, which significantly reduces the dimension of the output that needs to be learned.

##### B. Simulation Verification

Note that the end-to-end GAN accepts passband semantics  $c$  and random noise  $z$  but directly outputs the physical parameters. Therefore, we model the FSS structure based on the output of the GAN and monitor the corresponding  $S_{11}$  curve in the simulation. In this section, we will mainly verify the effectiveness of the end-to-end GAN from two aspects: precise controllability and generating diversity.

When the GAN is fed with a fixed random noise  $z$  and a gradient passband semantics  $c$ , the change of the  $S_{11}$  curve corresponding to the physical parameters given by the GAN is shown in Fig. 9. The three verification cases strongly illustrate that GAN successfully learns the mapping relationship of

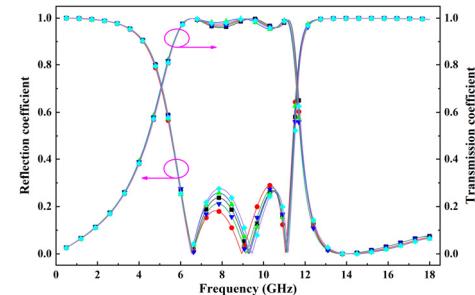


Fig. 10. Validation of GAN generative curve diversity. GAN accepts a fixed passband semantic  $c$  and different noises  $z$ , and generates diverse curves.

TABLE II  
COMPARISONS WITH FORWARD OPTIMIZATION METHOD

Method	Data Collection	Training Time	Execution Time	Iteration for new demands
GA Optimizer (CST)	/	/	~20h	✓
GA with forward NN	~2d	~6min	~10min	✓
The proposed GAN	~2d	~10min	~0.004s	✗

$\{R, H\} \rightarrow \{P\}$ , and the output physical parameters can be an effective solution to guide the design of the FSS passband. Fig. 10 further shows that the GAN as a generative model has the ability to give a variety of feasible solutions for a specified passband demand, while the discriminative model does not have one. This is a very innovative feature in the FSS inverse design, which enables engineers to select the appropriate solution for practical applications.

#### V. METHOD COMPARISONS AND DISCUSSION

In this section, the proposed method is briefly discussed and compared with other already existing methods. Table II shows the efficiency and advantages of the proposed method over different forward optimization algorithms. It should be noted that the first method used for comparison is the genetic algorithm (GA) optimizer embedded in the CST platform. The second one is a typical optimization method with the help of forward networks [16]. The third one is our proposed end-to-end paradigm GAN. To obtain a fair comparison, all methods are run on the same computer with NVIDIA GeForce RTX 3080 graphic processing unit (GPU).

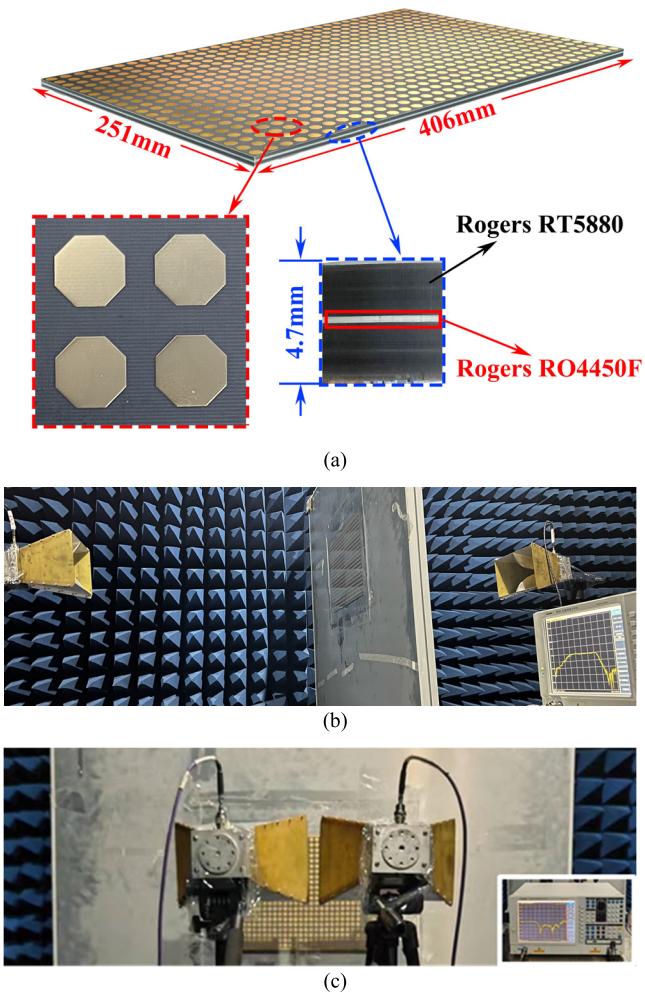


Fig. 11. (a) Fabricated sample. Measurement setup for (b) transmission coefficient and (c) reflection coefficient.

The primary efficiency of our proposed method is that the feasible solution to meet the design demands can be delivered immediately without time-consuming iterations. Although it took about two days to collect the training data, this overhead was considered a one-time event [17], [22]. Our efficiency advantage is further amplified when faced with multiple performance demands. The nonunique mapping problem in the direct inverse modeling method is efficiently solved by deploying a novel generative model. Since no additional custom training methods are required (e.g., data processing [18], tuning the dimensionality of network inputs [20], and outputs [21]), our approach is lighter and easily scalable. This proposed approach is considered more valuable for guiding engineering design because of its ability to accept a more realistic design requirement and provide multiple solutions.

## VI. EXPERIMENTAL VERIFICATION

A detailed case of designing an FSS structure with a passband of [6.0, 11.0] is shown here. Specifically, [6.0, 11.0] is fed to the well-trained end-to-end GAN as passband semantics  $c$ . Then, a feasible solution ( $d = 0.68012$ ,  $s = 2.9195$ ,  $h = 2.3459$ , and  $r = 4.2273$ ) of the physical parameters

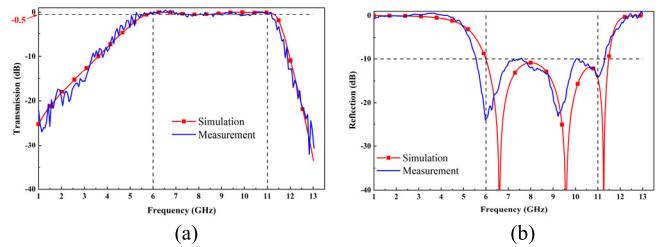


Fig. 12. Comparisons between the measured and simulated (a) transmission coefficient and (b) reflection coefficient. The red line represents the simulation results of the fabricated model and the blue line is the measurement. The simulation results are carried out by commercial simulation software CST.

is obtained. Considering the real manufacturing process, the corresponding FSS structure is fabricated according to the structure parameter ( $d = 0.7$ ,  $s = 2.9$ ,  $h = 2.3$ , and  $r = 4.2$ ). As shown in Fig. 11(a), the fabricated sample consists of  $34 \times 21$  units with an overall size of  $406 \times 251$  mm.

To demonstrate that the proposed method can effectively guide the inverse design, an FSS was manufactured according to the above parameters. The measurement was carried out in an anechoic chamber, as shown in Fig. 11(b). The instruments we used are listed as follows: a pair of horn antennas HD-10180 with frequencies ranging from 1 to 18 GHz and the vector network analyzer (VNA) 3656D (made by the 41st Research Institute of China Electronics Technology Group Corporation 41) covering the bandwidth from 0.3 MHz to 20 GHz. As shown in Fig. 12, the comparison between the simulated and measured S-parameters under normal incidence is given. It is concluded that the measured and simulated results achieve a good agreement, and the feasibility of the method proposed in this article is further verified in experiments.

## VII. CONCLUSION

In this article, we discuss the inverse design dilemma caused by the data nonuniqueness problem and propose deploying generative models as a solution. To design the FSS structure with a desired electromagnetic response, two approaches for deploying GANs are provided. The main idea of the first approach is to fully use the trained inverse network, which receives the electromagnetic response described in higher dimensions (such as the  $S_{11}$  curve) as input. Conditional generative models cGANs are deployed in this approach, which, on the one hand, receive the descriptive semantics of the desired electromagnetic response and, on the other hand, enable the diversity of the generated data through random noise variables. A neural network layer with inverse discrete-time Fourier transform function is proposed and employed to help the generative task of GAN. In another approach, we implement an end-to-end GAN that directly generates the physical parameters that guide the design of the FSS without relying on a trained inverse network. The effectiveness of all proposed methods has been closed-loop verified in simulation. An additional experimental test was carried out to confirm the feasibility of our method. A favorable agreement between simulation and measurement is achieved. Unlike designing or improving discriminative models, applying generative models

as a solution to the inverse modeling dilemma stands in a new perspective, which also paves the way for complex inverse design tasks for microwave components.

## REFERENCES

- [1] B. A. Munk, *Frequency Selective Surfaces: Theory and Design*. Hoboken, NJ, USA: Wiley, 2000.
- [2] B. A. Munk, *Finite Antenna Arrays and FSS*. Hoboken, NJ, USA: Wiley, 2003.
- [3] D. Li et al., "A low-profile broadband bandpass frequency selective surface with two rapid band edges for 5G near-field applications," *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 2, pp. 670–676, Apr. 2017.
- [4] D. Li, Z. Shen, and E.-P. Li, "Spurious-free dual-band bandpass frequency-selective surfaces with large band ratio," *IEEE Trans. Antennas Propag.*, vol. 67, no. 2, pp. 1065–1072, Feb. 2019.
- [5] R. Dubrovka, J. Vazquez, C. Parini, and D. Moore, "Modal decomposition equivalent circuit method application for dual frequency FSS analysis," in *Proc. Loughborough Antennas Propag. Conf.*, Apr. 2007, pp. 257–260.
- [6] T. Li et al., "A novel miniaturized strong-coupled FSS structure with excellent angular stability," *IEEE Trans. Electromagn. Compat.*, vol. 63, no. 1, pp. 38–45, Feb. 2021.
- [7] Z. Ma, K. Xu, R. Song, C. Wang, and X. Cheng, "Learning-based fast electromagnetic scattering solver through generative adversarial network," *IEEE Trans. Antennas Propag.*, vol. 69, no. 4, pp. 2194–2208, Apr. 2021.
- [8] X. Chen, Z. Wei, M. Li, and P. Rocca, "A review of deep learning approaches for inverse scattering problems (invited review)," *Prog. Electromagn. Res.*, vol. 167, pp. 67–81, 2020.
- [9] 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.
- [10] 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.
- [11] Q.-J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*. Norwood, MA, USA: Artech House, 2000.
- [12] R. Zhu et al., "Multiplexing the aperture of a metasurface: Inverse design via deep-learning-forward genetic algorithm," *J. Phys. D, Appl. Phys.*, vol. 53, no. 45, Nov. 2020, Art. no. 455002.
- [13] O. U. Rehman, S. U. Rehman, S. Tu, S. Khan, M. Waqas, and S. Yang, "A quantum particle swarm optimization method with fitness selection methodology for electromagnetic inverse problems," *IEEE Access*, vol. 6, pp. 63155–63163, 2018.
- [14] J. Peurifoy et al., "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.*, vol. 4, no. 6, Jun. 2018, Art. no. eaar4206.
- [15] Z. Zhang, B. Liu, Y. Yu, and Q. S. Cheng, "A microwave filter yield optimization method based on off-line surrogate model-assisted evolutionary algorithm," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 6, pp. 2925–2934, Jun. 2022.
- [16] Z. Gu, D. Li, Y. Fan, L. Zhang, Y. Liu, and E. Li, "Intelligent design of arbitrary bandstop FSS through deep learning and genetic algorithm," in *Proc. Asia-Pacific Int. Symp. Electromagn. Compat. (APEMC)*, Beijing, China, Sep. 2022, pp. 649–651.
- [17] Y. Jia et al., "In situ customized illusion enabled by global metasurface reconstruction," *Adv. Funct. Mater.*, vol. 32, no. 19, May 2022, Art. no. 2109331.
- [18] 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.
- [19] 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.
- [20] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," in *Proc. Conf. Lasers Electro-Opt.*, 2019, pp. 1–2.
- [21] 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.
- [22] E. Zhu, Z. Wei, X. Xu, and W.-Y. Yin, "Fourier subspace-based deep learning method for inverse design of frequency selective surface," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5130–5143, Jul. 2022.
- [23] I. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [24] M. Mirza and O. Simon, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.
- [25] W. Chen, K. Chiu, and M. Fuge, "Airfoil design parameterization and optimization using Bezier generative adversarial networks," 2020, *arXiv:2006.12496*.

**Zheming Gu** received the B.S. degree in communication engineering from the School of Mechanical, Electrical and Information Engineering, Shandong University, Jinan, China, in 2019. He is currently pursuing the Ph.D. degree in electronics science and technology with Zhejiang University, Hangzhou, China.



His current research interests include machine learning techniques for electromagnetic interference (EMI) analysis and optimization design of systems in package.

**Da Li** (Member, IEEE) received the B.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 2014 and 2019, respectively, both in electrical engineering.

From 2017 to 2018, he worked at Nanyang Technological University, Singapore, as a Visiting Scientist. From 2019 to 2021, he was with the Science and Technology on Antenna and Microwave Laboratory, Nanjing, China, as a Research Fellow. He is currently an Assistant Professor at Zhejiang University. His research interests include machine learning, antennas, metasurfaces, and electromagnetic compatibility. He has authored or coauthored more than 50 refereed papers and served as a reviewer for six technical journal articles.

Dr. Li is a TPC member of various IEEE conferences. He was a recipient of the Outstanding Young Scientist Award from the Asia-Pacific International Symposium on Electromagnetic Compatibility in 2022.



**Yunlong Wu** received the B.S. degree in systems engineering from Air Force Engineering University, Xi'an, China, in 2012. He is currently pursuing the M.S. degree in electronic science and technology with Zhejiang University, Hangzhou, China.

His current research interests include artificial electromagnetic structures and the application of machine learning in the field of electromagnetism.



**Yudi Fan** received the B.Eng. degree in information and communication engineering from the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, in 2019, where he is currently pursuing the Ph.D. degree.

From 2022 to 2023, he worked at the Singapore University of Technology and Design, Singapore, as a Visiting Researcher. His current research interests include microwave absorbers, reconfigurable intelligent surfaces, and active metasurfaces.



**Chengting Yu** received the B.S. degree from Zhejiang University, Hangzhou, China, in 2021, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering.

His current research interests include machine learning and neuromorphic computing.



**Hongsheng Chen** (Fellow, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Zhejiang University (ZJU), Hangzhou, China, in 2000 and 2005, respectively.

In 2005, he became an Assistant Professor with ZJU, where he was an Associate Professor in 2007 and a Full Professor in 2011. He was a Visiting Scientist from 2006 to 2008 and a Visiting Professor from 2013 to 2014 with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA. He is currently a Chang Jiang Scholar Distinguished Professor with the Electromagnetics Academy, ZJU. He has coauthored more than 200 international refereed journal articles. His works have been highlighted by many scientific magazines and public media, including *Nature Scientific American MIT Technology Review*, and *Physorg*. His current research interests include metamaterials, invisibility cloaking, transformation optics, graphene, and theoretical and numerical methods of electromagnetics. His research work on an invisibility cloak was selected in Science Development Report as one of the representative achievements of Chinese Scientists in 2007.

Dr. Chen serves as a regular reviewer for many international journals on electromagnetics, physics, optics, and electrical engineering. He serves as a Topical Editor for the *Journal of Optics* and the Editorial Board for Nature's *Scientific Reports and Progress in Electromagnetics Research*. He was a recipient of the National Excellent Doctoral Dissertation Award in China in 2008, the Zhejiang Provincial Outstanding Youth Foundation in 2008, the National Youth Top-Notch Talent Support Program in China in 2012, the New Century Excellent Talents in University of China in 2012, the National Science Foundation for Excellent Young Scholars of China in 2013, and the National Science Foundation for Distinguished Young Scholars of China in 2016.



**Er-Ping Li** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Sheffield Hallam University, Sheffield, U.K., in 1992.

From 1989 to 1992, he was a Research Associate/Fellow with the School of Electronic and Information Technology, Sheffield Hallam University. From 1993 to 1999, he was a Senior Research Fellow, a Principal Research Engineer, and the Technical Director of the Singapore Research Institute and Industry, Singapore. Since 2000, he has been serving as the Director for the Advanced Electronic Systems

and Electromagnetics Department and the Senior Director for collaborative research with the Singapore National Research Institute of High-Performance Computing, Singapore. Since 2010, he has been with Zhejiang University, Hangzhou, China, where he is currently a Changjiang-Qianren Distinguished Professor in radio frequency and nanoelectronic research and the Dean of the Zhejiang University/University of Illinois at Urbana-Champaign Institute. He has published more than 400 papers in the referred international journals and conferences and authored two books published by Wiley/IEEE Press and Cambridge University Press. His research interests include fast and efficient computational electromagnetics, microscale/nanoscale integrated circuits and electronic package, electromagnetic compatibility (EMC), SI, and nanotechnology.

Dr. Li is a fellow of the Electromagnetics Academy, USA. He was an elected Distinguished Lecturer of *IEEE Electromagnetic Compatibility Magazine* (EMC) from 2007 to 2008. He received numerous awards, including the *IEEE Electromagnetic Compatibility Magazine* (EMC) Technical Achievement Award and the IEEE Richard Stoddard Award in EMC. He was the President of the International Zurich Symposium on EMC, Singapore, in 2006 and 2008; the General Chair of the 2008, 2010, and 2012 Asia-Pacific International EMC Symposium; and the Chairperson of the *IEEE Electromagnetic Compatibility Magazine* (EMC) Singapore Chapter from 2005 to 2006. He has been an invited speaker for numerous talks and keynote speeches at various international conferences and forums. He is an Associate Editor of *IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY* and *IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY*.