

# Tandem Neural Network Based Design of Multiband Antennas

Aggraj Gupta<sup>ID</sup>, Student Member, IEEE, Emir Ali Karahan<sup>ID</sup>, Graduate Student Member, IEEE, Chandan Bhat<sup>ID</sup>, Kaushik Sengupta<sup>ID</sup>, Senior Member, IEEE, and Uday K. Khankhoje<sup>ID</sup>, Senior Member, IEEE

**Abstract**—We present a deep neural network-based framework for designing multiband microstrip antennas given a desired impedance matching spectrum. The approach enables a design methodology that generates the desired antenna structures rapidly (under a second) through an effective deep learning-enabled search of a large design space and eliminates the need for extensive domain knowledge of antenna design. The framework is built on our innovations in tandem neural networks consisting of two cascaded neural networks. Our structures are parameterized in an exponentially large design space of discrete variables (pixels), leading to the realization of nonintuitive structures. This end-to-end synthesis in terms of discrete variables is enabled by introducing a new type of “smooth thresholding” (ST) activation function, which, along with crucial regularization terms in the network loss function, aids in designing our structures. We perform extensive neural network optimizations and study various trade-offs in the design process. We demonstrate the efficacy of our methods by generating single and dual-band resonant structures, which can be up to 50% more compact in terms of area, and up to 18% thinner in terms of substrate height than conventional structures, while retaining competitive performance parameters in terms of gain, polarization properties, radiation efficiency (RE), and fractional bandwidth (FBW).

**Index Terms**—Antennas, artificial intelligence, design automation, microstrip antennas, optimization methods.

## I. INTRODUCTION

IN RECENT years, ideas from machine learning have started to play a pivotal role in the design of electromagnetic and photonic structures [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. We revise the historical context of antenna design in order to better appreciate the ideas that follow: For decades, the paradigm of antenna design has been characterized by two features, one, expensive electromagnetic simulations to predict the performance of an antenna, and two, reliance on vast domain knowledge and rules of thumb to fine-tune a design. In a sense, the second feature is an outcome of the first because

Manuscript received 19 December 2022; revised 24 March 2023; accepted 2 May 2023. Date of publication 19 May 2023; date of current version 4 August 2023. This work was supported by the Ministry of Electronics and Information Technology (MeitY), Government of India, through the Project “6G: Sub-THz Wireless Communication with Intelligent Reflecting Surfaces (IRS)” under Grant R-23011/3/2022-CC&BT-MeitY. (Corresponding author: Uday K. Khankhoje.)

Aggraj Gupta, Chandan Bhat, and Uday K. Khankhoje are with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai 600036, India (e-mail: uday@ee.iitm.ac.in).

Emir Ali Karahan and Kaushik Sengupta are with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: kaushiks@princeton.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAP.2023.3276524>.

Digital Object Identifier 10.1109/TAP.2023.3276524

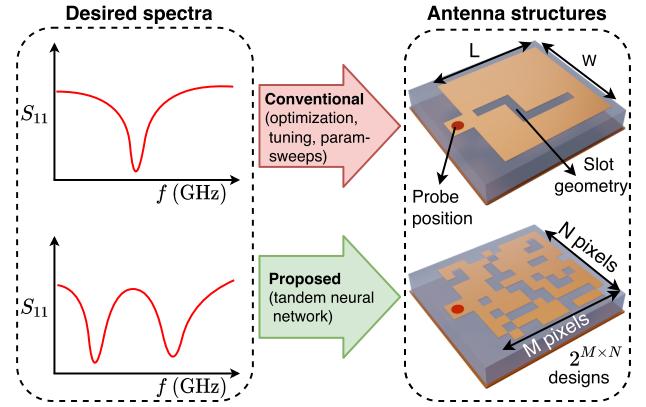


Fig. 1. Overview of the conventional and proposed framework of antenna design. Traditionally, a single-band antenna can be designed by varying a finite set of parameters (such as length, width, probe position, and slot geometry) in the microstrip patch. Designing a multiband antenna is much more challenging. The proposed framework uses a “tandem” neural network to design an antenna with a given spectrum in less than a second.

numerical optimization for antenna design is extremely challenging in the face of time and memory-intensive electromagnetic simulations. However, with the advent of fast computing hardware and the phenomenal success of machine learning in diverse fields from image processing to economic forecasting, the field of electromagnetic and photonic device design too stands to undergo a paradigm shift [11], [12].

In this contribution, we propose a way of thinking that revises both aspects of the traditional antenna design mentioned above. As a result, we have an end-to-end machine learning paradigm that generates an antenna given a desired performance as input. This effectively cuts out the reliance on domain knowledge; in fact, existing domain knowledge can be incorporated into the training datasets of the new approach. Doing so allows us to enlarge the possibilities of antenna performance that we could ask for and paves the way for *multifunction* devices [10]. For example, can we think of designing an antenna resonant at two bands, with independent control over the beam angle and polarization of each band? It would be difficult to answer this question in the affirmative with conventional trial and error methods, but such an antenna is a very likely possibility in the new paradigm using machine learning. As an illustration, we draw the reader’s attention to Fig. 1 which discusses the design of a single- and dual-band antenna. While designing a single-band antenna using conventional methods is a matter of simply varying the dimensions, the story gets complicated when we go to dual-band antennas. Traditionally, this has been

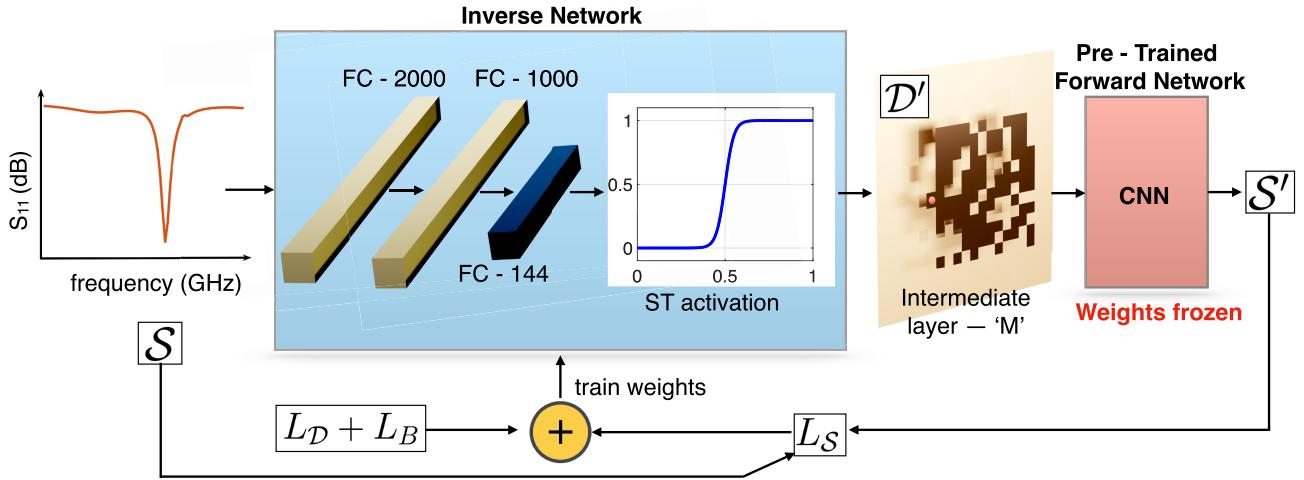


Fig. 2. Proposed tandem network architecture for antenna design featuring a set of layers that take antenna performance,  $\mathcal{S}$ , as input and produce the design of a tessellated patch antenna at an intermediate layer,  $\mathcal{D}'$ ; this design is cascaded with the pretrained surrogate model for electromagnetic simulations that takes  $\mathcal{D}'$  as input and produces its response  $\mathcal{S}'$  as output. A novel ‘ST’ activation layer helps produce a binary antenna design. The network loss functions  $L_D$ ,  $L_S$ ,  $L_B$  are elaborated in Section III.

accomplished by introducing parasitic elements, slots, notches, lumped elements, shorting posts, fractal elements, etc., and the parameters varied using trial and error till the desired antenna is achieved (for e.g., [13], [14])—all in all, a time-consuming process heavily reliant on prior experience and domain knowledge. On the other hand, with the techniques introduced in this work, we can design either kind of antenna in less than a second (on even an ordinary laptop) in a fully automated manner by using designed trained neural networks.

#### A. Overview of Our Approach

In this work we demonstrate the design of antenna structures by developing a “tandem” neural network, as seen in Fig. 2. This network consists of an inverse network and a deep convolutional neural network (CNN) that acts as an ultrafast surrogate for expensive electromagnetic simulations. The operation of two cascaded neural networks is what earns it the name ‘tandem’, first introduced in the context of photonic structures [15]. The purpose of the cascade is to solve the data inconsistency problem, which arises because there can be multiple structures that give nearly the same performance. As a result, a naive attempt to train a neural network directly between the device performance (input) and the design (output) fails to converge; the cascade solves this problem by computing the response of the generated design, which leads to a consistent objective function for the purpose of training [15]. The device performance can be characterized by the antenna return loss ( $S_{11}$  as a function of frequency) as well as other features such as the radiation pattern and polarization. We describe an antenna by a variable ‘chessboard’ pattern of metallic sub patches (see Fig. 3 for e.g.,) thereby giving us an exponentially large design space [16]. This larger space allows us to explore antenna designs that would otherwise be missed in an approach that uses a templated geometry [10].

#### B. Related Work

Recent work reported the use of neural networks for multiobjective antenna design based on templated geometries

[9], [10], [17], [18], [19]. Characterizing antenna design by means of template parameters [20] economizes the number of variables, thus leading to simpler neural networks and lesser data requirements, but at the cost of a smaller design space. Earlier work in nanophotonic structure design has shown how a surrogate neural network for the forward electromagnetic simulations can be re-purposed for the task of inverse design by using the backpropagation algorithm [6]. This is convenient to do when the design variables are continuous-valued, since the backpropagation algorithm requires differentiability of the variables. However, when the design variables are discrete valued the task becomes harder and approaches such as level sets [21] and evolutionary algorithms [3], [16], [22], [23] are common in this case. In [24] a deep CNN-based model is trained to learn the resonant frequencies of dual-band pixelated antennas. While this approach only learns the resonant frequencies, we take a more general approach to learning the  $S_{11}$  response in a frequency range of interest. Indeed, in [25], we have innovated the coupling of a deep CNN used as a surrogate forward solver, with genetic algorithm and binary particle swarm optimization (BPSO) methods for the synthesis approach, leading to compact, multiband antennas.

Moving on from conventional algorithms, a promising all-neural network approach for inverse design compatible with both continuous and discrete valued design variables involves the use of tandem neural networks. The first tandem network [15] designed a one-dimensional stack of variable-thickness dielectric layers with a target transmission spectrum using an objective function consisting of the spectrum loss,  $\mathcal{S} - \mathcal{S}'$  (see Fig. 2), where  $\mathcal{S}, \mathcal{S}'$  refer to the desired and surrogate predicted spectra, respectively. A subsequent improvement was achieved [26] by adding a ‘design’ loss to the objective function, i.e.,  $\mathcal{D} - \mathcal{D}'$ , where  $\mathcal{D}, \mathcal{D}'$  refer to a design and its prediction via the inverse network, respectively; doing so led to more accurate device designs. Recently, the tandem approach was used to design wideband Schiffman phase shifters [27] by minimizing a combination of phase error and return loss during network training. While the previous

three works dealt with continuously valued design variables, a subsequent work offered a route toward designing structures with discrete parameters [28]. The innovation here was to include a term of the type ‘ $d(d - 1)$ ’ into the objective function ( $d$  is a design variable) which promoted solutions to take either of the discrete values  $d = \{0, 1\}$ . Finally, note that while the tandem architecture superficially resembles that of a conventional autoencoder (CAE) [29] in that both the input and output refer to the same quantity, the significant difference is that the CAE learns a coded representation of the input whereas the tandem approach specifies the equivalent representation to be the device geometry.

### C. Our Contributions

The present work builds on all the previous tandem network-related works [15], [26], [28] and further introduces a new type of activation function which we term as “ST”, short for Smooth approximation of a Thresholding function, to better promote the discrete nature of the design. Through extensive numerical simulations, we demonstrate the superior nature of this innovation in tandem networks, leading to the rapid design of miniaturized single- and dual-band antennas. We explore the trade-offs faced in the architecture of these neural networks, particularly in relation to the discrete nature of the design variables.

The article is organized as follows. In Section II, we describe the neural networks developed in this work, as well as go over details of the datasets required for network training. After describing training details for both, the forward and tandem network, we elaborate on our innovations in network design and hyperparameter choice which lead to optimal performance. In Section III, we report the various types of antennas that are generated by deploying the network, including single- and dual-band compact structures. In Section IV we discuss certain limitations of our approach while simultaneously proposing workarounds as well as future extensions that can be considered. Finally, we conclude in Section V.

## II. BUILDING BLOCKS FOR NEURAL-NETWORK BASED INVERSE DESIGN

Here, we describe the overall framework of the proposed approach for antenna design using a tandem neural network, starting with a description of the network details, followed by training specifications, and finally a summary of the innovations in our approach.

### A. Tandem Neural Network

**1) Antenna Structure:** To enable convergence toward synthesis of arbitrary-shaped planar structures with the desired radiation properties, we first discretize the space to a moderately large number of pixels, that can still approximate functionalities achievable by continuously shaped antennas [Fig. 3(a) and (c)]. As described earlier, a candidate antenna structure is generated by starting with a conventional metallic patch antenna and tesselating the surface into a  $12 \times 12$  sub-patches. Each of these sub-patches is characterized by a “1,”

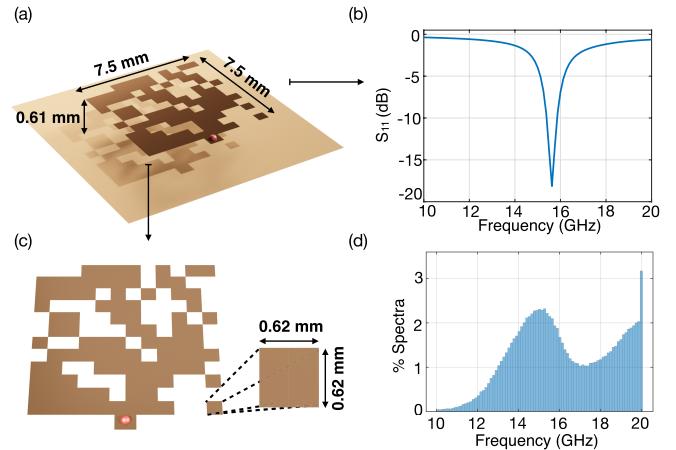


Fig. 3. (a) Sample of an air filled pixelated microstrip antenna with a coaxial feed. The gap between the upper and ground plate is 0.61 mm and the thickness of metal used is 0.035 mm. The inner conductor of the coaxial connector extends through the air and is connected to the patch, while the outer conductor is connected to the ground plane. (b) Return loss ( $S_{11}$  in dB). (c) Top view of the antenna without the ground plane (red dot is the feed location). (d) Histogram of resonance spectra (in percentage) (180 k out of 500 k dataset samples have a resonance).

meaning metal, or “0,” meaning no metal. Thus, an antenna structure is parameterized by 144 discrete variables, giving us a large exponential design space, typically much larger than what can be effectively searched by evolutionary algorithms (such as a genetic algorithm, BPSO etc.).

As mentioned in Section I, the tandem architecture [15] avoids the data inconsistency problem that arises due to the nonunique correspondence between a device design and its spectrum by the cascade of two neural networks. Fig. 2 shows the schematic of the tandem network, which consists of an inverse network connected to a pretrained forward network with frozen weights. The input to the tandem network is the desired spectrum ( $\mathcal{S}$ ), and the output is the predicted spectrum ( $\mathcal{S}'$ ). The inverse network (see Fig. 2) consists of three dense layers, which are each followed by batch normalization and leakyReLU activation functions. The output of the final leakyReLU activation layer is connected to a layer with the newly introduced ST activation for generating binary outputs. The ST activation function is shown in the inset of Fig. 2 and is defined as

$$f(x) = \frac{1}{2} + \frac{1}{2} \tanh \left[ m \left( x - \frac{1}{2} \right) \right] \quad (1)$$

where  $m$  is the thresholding hyperparameter. The ST activation function acts as follows: if  $x \leq 0.5$ , then  $f(x) \leq x$  (damps); if  $x \geq 0.5$ , then  $f(x) \geq x$  (amplifies), and this rate of damping/amplification is governed by  $m$ . We note that ST activation function is a generalization of the unipolar sigmoid function with  $m = 1$  [30].

**2) Forward Surrogate Model:** The forward surrogate model approximates the antenna performance, in particular the antenna return loss given the antenna design. This surrogate is itself achieved via a deep CNN, which treats the input as a single channel image, and through a sequence of convolutional filters and fully connected layers, gives the desired response

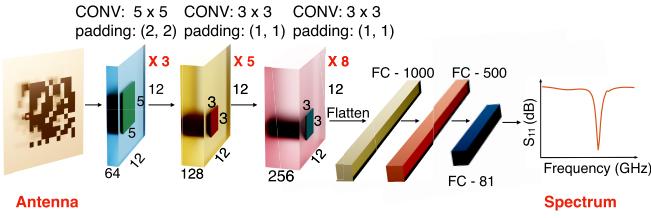


Fig. 4. Deep CNN forward surrogate model consisting convolutional and fully connected layers, denoted by “CONV” and “FC,” respectively. A tessellated antenna structure is the input and the antenna return loss is the output.

as the output [25]. Once the CNN is trained, it approximates a complex electromagnetic simulation (such as integral equation or finite-difference time-domain simulations) in orders of magnitude lesser time [6]. The network contains 56 layers with filters, weights, and biases; the first 16 combine convolutional 2-D, batch normalization, and the leaky ReLU activation function layers, as shown in Fig. 4. The next two are fully connected layers; batch normalization, leakyReLU activation function, and dropout (value = 0.4) layers are applied to the output of every fully-connected layer. The output of the last fully-connected layer is fed to an 81-D output regression layer.

### B. Network Training

1) *Dataset Generation:* To generate a training dataset, we start with a “mother” metallic patch of size  $7.5 \times 7.5$  mm and tessellate it into  $12 \times 12$  pixels. This structure has a weak resonance near 19.5 GHz. The pixels values of 0 or 1 are populated randomly, while ensuring that the two pixels adjacent to the feed location are always metal to ensure connectivity. This gives us an exponentially large design space of nearly  $2^{142}$  possible structures. The return loss corresponding to the antenna is simulated at 81 equispaced points in a 10–20 GHz frequency range using the MATLAB Antenna Toolbox. A total of 500 k antennas and their spectra are generated, which are distributed in a 80:10:10 ratio for training, validation, and testing, respectively. We define a resonant structure to be one which displays a dip below  $-10$  dB in the return loss,  $S_{11}$ , at one or more points in the frequency range. We find that most of the 500 k samples have no resonances, while about 180 k are resonant structures; further, we plot a histogram of the resonant frequencies from the dataset in Fig. 3(d). The histogram reveals an “unbalanced” dataset, in the sense that the resonances are not uniformly distributed over the frequency range.

We make an important remark about a design choice in the dataset generation. Considering the large size of the dataset, it was necessary to keep the simulation time for a single antenna to be as low as possible. We choose to have air rather than dielectric as the substrate for the antenna structure as this considerably lowers the simulation time (by a significant factor of 50–60). Parallelizing the dataset generation over 400 cores of the high-performance computing facilities at the Princeton Research Computing Resources center (consisting of a mix of Intel Ice Lake and Cascade Lake processors), the dataset

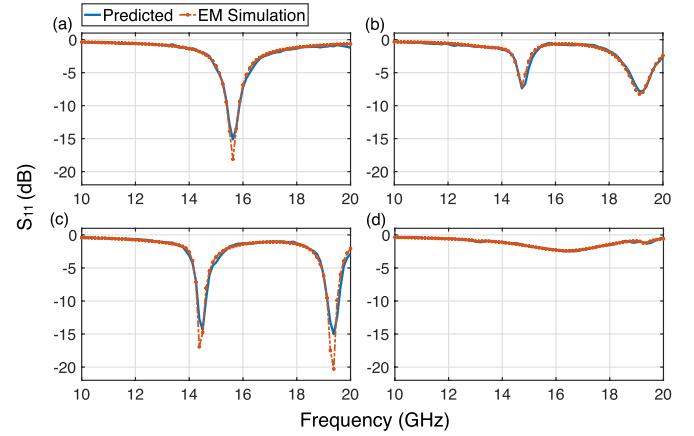


Fig. 5. Comparison of the predicted  $S_{11}$  spectrum from the trained forward model with the true spectrum from electromagnetic simulator showing the efficacy of the forward model: (a) and (c) resonant structures (b) and (d) nonresonant structures.

was generated in approximately 24 h. Fig. 3 shows a sample antenna and its  $S_{11}$  spectrum.

We note that this choice of substrate is a matter of convenience and does not imply any fundamental limitation on our *approach*. Since Maxwell’s equations are scale-invariant, we can scale these antenna structures to different operating frequencies by performing an overall scaling of the dimensions. However, scale invariance can not be used to infer the response of a dielectric-filled antenna given that of an air-filled one. To this end, in [25] (and further elaborated in Section IV-B), we show how a transfer learning approach enables the dataset of the air-filled structures to be repurposed to rapidly learn the  $S_{11}$  response of dielectric-filled structures.

2) *Forward Network Training:* The loss function that characterizes the forward model is a mean squared error function of the form

$$L_f = \frac{1}{P} \sum_{i=1}^P \text{MSE}(\mathcal{S}_i, \mathcal{S}'_i) \quad (2)$$

where  $P$  is the batch size, set to 256. We choose NAdam [31] optimizer to update the weights and biases of the CNN. It can be seen from Fig. 5 that the forward network predicts the return loss of an antenna from the test dataset to a satisfactory level and can be substituted for an electromagnetic simulator. The forward network is trained in PyTorch using Google Colab GPU services for 15 epochs taking approximately 3 h. The training and validation loss are 0.62 and 0.71, respectively.

3) *Tandem Network Training:* The overall loss function,  $L_I$ , that characterizes the tandem network is given as

$$L_I = L_S + \alpha L_D + \beta L_B \quad (3)$$

where the constituent loss functions are

$$L_S = \text{MSE}(\mathcal{S}, \mathcal{S}') \quad (\text{spectrum loss}) \quad (4a)$$

$$L_D = \text{MSE}(\mathcal{D}, \mathcal{D}') \quad (\text{design loss}) \quad (4b)$$

$$L_B = \frac{1}{N} \sum_{i=1}^N (d'_i(d'_i - 1))^2 \quad (\text{binary loss}) \quad (4c)$$

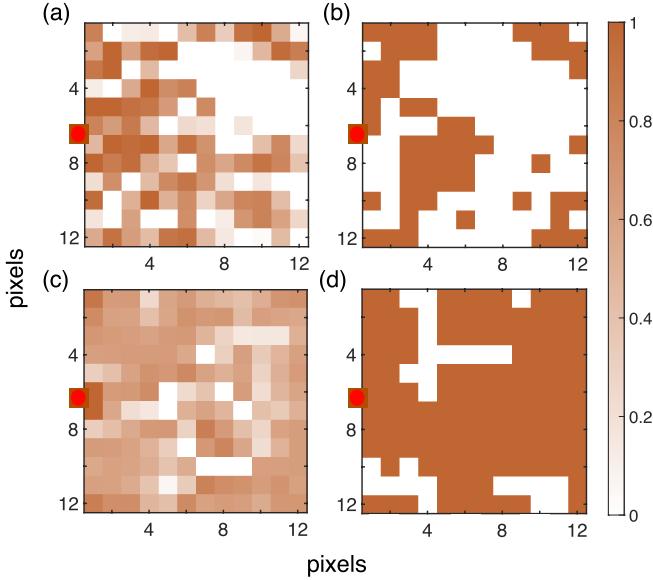


Fig. 6. Predicted antenna design with different loss functions: (a)  $L_I = L_S + 5L_B$ , (b)  $L_I = L_S$ , (c)  $L_I = L_S + 1L_D$ , and (d)  $L_I = L_S + 17L_D + 1L_B$ . The networks are trained (a) without ST activation layer; (b)–(d) with ST activation layer.

where,  $N$  is the total number of pixels,  $\alpha$  and  $\beta$  are the tuneable hyperparameters;  $\mathcal{S}'$  is the predicted spectrum,  $\mathcal{D}'$  is the predicted antenna design obtained in the intermediate layer ( $M$ ) in the tandem network,  $d'_i$  corresponds to the  $i$ th pixel in  $\mathcal{D}'$ , and MSE stands for mean squared error. Compared to prior work on nanophotonic structures [15], [26], [28], this is the first work to the best of our knowledge to combine all three loss functions along with the newly proposed ST activation to allow deep learning based inverse synthesis of antennas.

The network is trained with a batch size of 512 for 300 epochs taking approximately 12 h. The weights are updated using RAdam optimizer [32] and the hyperparameters  $\alpha$ ,  $\beta$ ,  $m$  are set to 17, 1, and 20, respectively, after an extensive grid search. The split-up of the loss function  $L_I$  into its constituents [as per (3)] in the format  $\{L_S, L_D, L_B\}$  is  $\{1.45, 6.45, 7 \times 10^{-7}\}$  (training) and  $\{1.48, 6.55, 8 \times 10^{-7}\}$  (validation) at the end of 300 epochs. The optimizers for the forward and tandem networks are chosen empirically by evaluating the training and validation loss over various solvers such as the steepest gradient (sgd), Adam, NAdam, and RAdam optimizers.

### C. Innovations in the Tandem Network

Having presented the various details of the tandem network and associated training details, we now highlight the rationale behind the design choices in conceiving the network.

**1) Role of ST Activation in Generating Binary Designs:** Prior to this work focusing on the design of nanophotonic structures [28] has shown techniques to deal with a small number of binary variables (6 compared to 142 in our case) in a tandem networks, where a binary design is obtained by training the network with spectrum *and* binary loss terms. In order to verify if adding a binary loss term is sufficient in our case, we implemented the tandem network without

the ST activation function, and a typical result is shown in Fig. 6(a) which shows the predicted design (after a sweep for the best value of hyperparameter  $\beta$ ); as can be seen, the obtained design is not binary. A possible solution to this problem is to simply threshold the obtained pixel values; for e.g., in [33] a metasurface is a  $64 \times 64$  binary image obtained by a thresholding operation. We find this approach to not be suitable, as the spectrum can change drastically as a result of thresholding. Hence it is necessary for the inverse design to give binary outputs, and we achieve this via the ST activation. As can be seen in Fig. 6(b) which is the result of using the loss function  $L_I = L_S$  along with an ST activation layer, the obtained design is binary. We note that while the design is binary, the obtained design is not practical due to the sparse distribution of metal patches and problems of feed connectivity. We address this issue next.

**2) Importance of Design and Binary Loss:** The design and binary loss terms, in addition to spectrum loss, are critical for the network to reach an optimum solution. Each term in the proposed loss function affects the design of the structures, as evident from Fig. 6: the regularized loss function with optimized hyperparameters promotes binary structures (due to  $L_B$ ) for which the return loss response is aligned with the desired spectrum (due to  $L_S$ ), and are suitable for the fabrication purpose (due to  $L_D$ ). We examine the importance of each term of the proposed loss function via the following experiments. First, we explore the role of having a spectrum loss along with a design loss, i.e.,  $L_I = L_S + \alpha L_D$ , on the lines of [26], plus the ST activation layer to promote a binary design. A resulting antenna can be seen in Fig. 6(c) which shows the design to be far from discrete (and therefore not practically realizable). In order to promote binary designs, we add a binary loss term to the loss function on the lines of [28], i.e.,  $L_I = L_S + \alpha L_D + \beta L_B$ ; the resulting antenna is seen in Fig. 6(d) which shows a binary design without any connectivity issues at the feed. The spectrum of the obtained antenna is independently simulated via commercial electromagnetics software (MATLAB) and found to be the same as predicted by the forward surrogate model. We thus conclude that in order to reliably get an accurate binary design it is essential for the network's loss function to have a design and binary loss terms in addition to the usual spectrum loss, and also an ST activation function.

**3) Choice of  $m$  in ST Activation:** The ST activation function produces binary outputs with high probabilities but fails when the input is in the close vicinity of 0.5 as can be inferred from (1). Fig. 6(c) showed that the design is nonbinary even with the ST activation included. The transition rate of the activation function from zero to one depends on the hyperparameter,  $m$ . Fig. 7 shows the spectrum and binary loss for  $m = 15$  and 20. Ideally, the binary loss should be zero for a purely binary design. At  $m = 20$  the binary loss is zero for epochs greater than 150. Though the spectrum loss is always lower for  $m = 15$ , the binary loss never reaches zero, and further, for epochs beyond the point 'K,' in Fig. 7, the binary loss increases. This is ostensibly happening because the network has moved into a local-minima region of nonbinary (and therefore *nonphysical*) designs which still show a lower

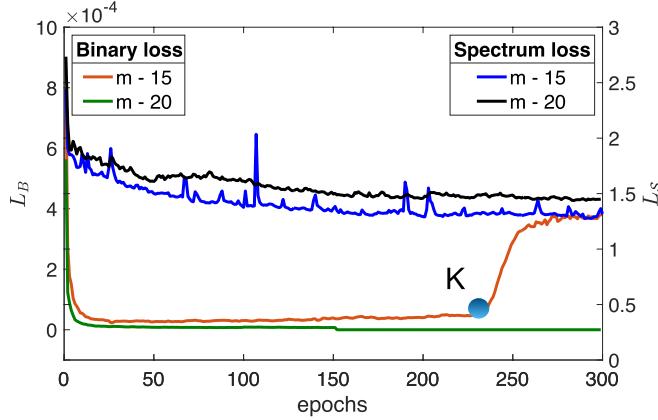


Fig. 7. Spectrum convergence and binary loss tradeoff for different  $m$  values in the training of the tandem network. The loss function consists of all three terms, i.e., spectrum, design, and binary loss terms. ‘ $K$ ’ is the trade-off point that helps us fix the hyperparameter  $m$ . At the end of 300 epochs, the design loss is 7.5 for  $m = 15$  and 6.6 for  $m = 20$  (not shown here).

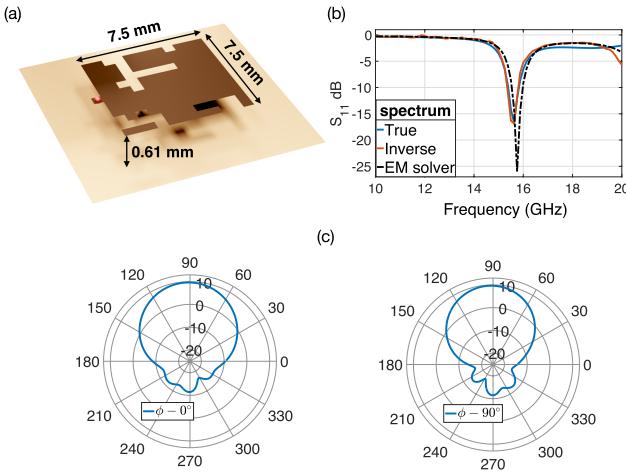


Fig. 8. (a) Inverse designed single-band antenna, (b) reconstructed spectrum, and (c) elevation radiation pattern ( $\phi = 0^\circ$  and  $\phi = 90^\circ$ ).

overall loss function. This trade-off between the spectrum and binary loss helps us to fix the ST activation hyperparameter at  $m = 20$ .

### III. INVERSE DESIGNED ANTENNA STRUCTURES

In this section, we report various structures designed by the tandem network, including single- and dual-band antennas. In the case of single-band antennas, we also make comparisons with conventional patch (CP) antennas to demonstrate a key result of our work, that of the compactness of the obtained designs. In all the examples shown below, the return loss has been validated against rigorous MATLAB-based electromagnetic simulations.

#### A. Simulation Results—Single Band Antennas

A single-band compact resonant antenna design is chosen as the first numerical example for the evaluation of the proposed approach. Fig. 8 shows the design of the compact

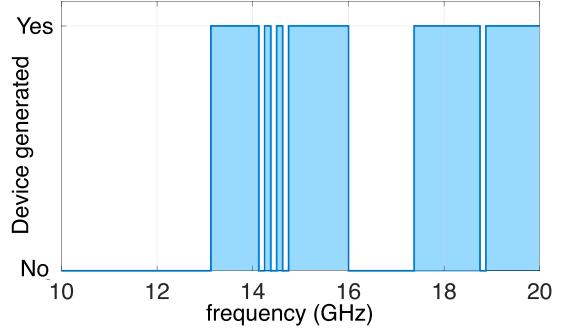


Fig. 9. Design range of predicted antennas for a single band spectrum from the inverse tandem network. “Yes” indicates the generated antenna has a single band spectrum in the frequency range shown in the shaded blue region while “No” means the tandem network could not predict a single band antenna.

TABLE I

DESIGN PARAMETERS FOR THE SINGLE-BAND RESONANT CP AND THE INVERSE DESIGNED PATCH (IDP) ANTENNA FOR DIFFERENT FREQUENCIES IN THE DATASET. CPL: CP LENGTH; CPW: CP WIDTH;  $f_0$ : RESONANT FREQUENCY,  $\Delta$  AREA: COMPACTNESS RELATIVE TO CP, RE, BW, FBW

CPL (mm)	CPW (mm)	$f_0$ (GHz)	Area ( $mm^2$ )		$\Delta$ area %	Gain (dB)		BW (FBW) MHz (%)	RE (%)
			(CP)	(IDP)		CP	IDP		
10.1	10.9	13.8	109	56.25	50	9.71	8.68	250 (1.8%)	91.2
9.85	10.7	14	105.4	56.25	48	9.71	9.45	250 (1.8%)	93.5
9.1	10	15	91	56.25	38	9.65	9.36	500 (3.33%)	93.2
7.48	8.32	18	62.23	56.25	10	9.68	9.29	500 (2.7%)	92

antenna (–31% smaller in area than a conventional antenna at the same frequency) at 15.8 GHz and its corresponding reconstructed and the simulated spectrum using the proposed tandem approach. It is seen that the spectrum of the inverse-designed antenna is in good agreement with the desired and electromagnetic-simulated spectrum. The axial ratio of the antenna in Fig. 8 is 56 (dB). Although the axial ratio and the radiation pattern of the inverse-designed antenna are not taken as a part of the cost function for optimization in the current work, the radiation pattern of the inverse-designed antenna is well-shaped and directed at  $\theta = 90^\circ$  (Gain: 9.635 dB). The device is found to be linearly polarized with the radiation efficiency (RE) above 90% considering Cu losses, which is comparable with conventional microstrip antennas [13].

Next, we compare several single-band antennas with their conventional counterparts in Table I. CP antennas cover an area of approximately  $(\lambda/2) \times (\lambda/2)$  [13], while in this work the antennas occupy an area on the order of  $(\lambda/4) \times (\lambda/4)$ , which is significantly compact at the design frequency.

We further check the rigor of our proposed tandem network on *never-seen-before* synthetic data. We choose a rectangular spectrum as a synthetic single-band  $S_{11}$  spectrum to check the reconstruction across the frequency span (10–20 GHz) using the proposed approach. Such a spectrum, with  $S_{11} = 0$  dB outside the passband and  $S_{11} = -25$  dB within a 500 MHz passband is not in the training dataset due to its obvious artificial nature. Fig. 9 shows the design range for the reconstruction of the single-band synthetic spectrum. The gaps in the plot are attributed to the imbalance of the single-

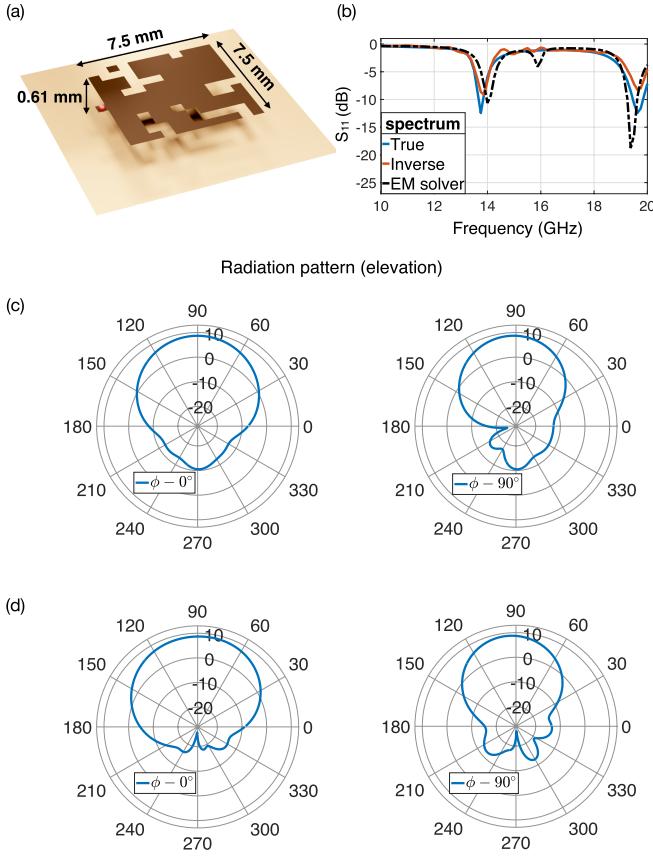


Fig. 10. (a) Inverse-designed dual-band antenna, (b) reconstructed spectrum, (c) elevation radiation pattern ( $\phi = 0^\circ$ ) of the inverse-designed antenna @  $f_1 = 14$  GHz (gain 9.73 dB), and (d) elevation radiation pattern ( $\phi = 90^\circ$ ) of the inverse-designed antenna @  $f_2 = 18.37$  GHz (gain 9.43 dB).

TABLE II

DUAL-BAND SPECTRA OF INVERSE DESIGNED ANTENNAS; HERE  $\lambda_0$  DENOTES THE WAVELENGTH CORRESPONDING TO THE FIRST FREQUENCY BAND,  $f_1$ , BW, FBW, RE

$f_1$ (GHz)	$f_2$ (GHz)	Gain (dB)		BW (FBW) MHz (%)		Aperture area	RE (%)	
		$f_1$	$f_2$	$f_1$	$f_2$		$f_1$	$f_2$
14.38	16.62	9.17	7.79	250 (1.7%)	250 (1.5%)	0.13 $\lambda_0^2$	90.6	88.1
14.5	17	9.01	7.88	250 (1.72%)	250 (1.47%)	0.13 $\lambda_0^2$	90.4	89.2
15.5	19.85	9.45	9.21	375 (2.41%)	375 (1.88%)	0.15 $\lambda_0^2$	91.2	90.5
14	18.37	9.73	9.43	250 (1.78%)	250 (1.36%)	0.12 $\lambda_0^2$	92.3	91.4

band resonant frequencies in the dataset. We observe that the blue region in Fig. 9 matches the high histogram regions of Fig. 3(d). This can be improved in the future by augmenting the training dataset to contain an equal number of samples for each simulated frequency.

### B. Simulation Results—Dual Band Antenna

Fig. 10 shows a dual-band inverse-designed antenna and the associated radiation patterns at the frequencies 14 and 18.37 GHz using the proposed tandem approach; the gain and the elevation radiation pattern are like that of a CP antenna. We expect to be able to have designer radiation patterns by incorporating the radiation pattern into the network loss function for optimization in the future.

TABLE III  
STATE-OF-THE-ART COMPARISON FOR DUAL-BAND MICROSTRIP ANTENNAS

Reference	$f_1$ (GHz)	$f_2$ (GHz)	Gain (dB)		Bandwidth (MHz)		Aperture area
			$f_1$	$f_2$	$f_1$	$f_2$	
[14]	2.4	5.6	9.55	8.50	97	360	$1.16 \lambda_0^2$
[34]	2.4	5.8	6.8	2.1	164	256	$0.64 \lambda_0^2$
[35]	1.57	2.45	7.47	7.07	47	125	$0.64 \lambda_0^2$
[36]	2.4	5.2	4.1	1.4	90	277	$0.08 \lambda_0^2$

Table II shows different dual-band designs, where we also report the area of obtained antennas in terms of the wavelength of the lower frequency band,  $\lambda_0$ .

We have also compared the run-time of deep learning-assisted evolutionary algorithms such as the modified BPSO [37] and the Genetic Algorithm (GA) [38] to generate the antenna structures corresponding to the spectra in Figs. 8(b) and 10(b); single band structures took 6–7 min, while double band structures took 10–12 min. In contrast, the proposed approach generated the structures nearly instantaneously.

### C. Comparison With Conventional Antennas

1) *Aperture Area:* We draw the reader's attention to Table III, where a comparison of various dual band structures is presented. In particular, we note that our designs yield much more compact antennas than those reported in the literature. For instance, the lowest area achieved by our work is  $0.12\lambda_0^2$ , while these are the areas reported in other works:  $1.16\lambda_0^2$  for a dual mode circular patch antenna in [14],  $0.64\lambda_0^2$  for an E-shaped microstrip patch in [34] and an arc-shaped slot patch antenna in [35], and  $0.08\lambda_0^2$  for a shorted microstrip antenna in [36]. It must be noted in the latter case that the lower aperture area is coming at the cost of significantly lower antenna gain as compared to our antennas. It is important to note that the structures designed by us have an air substrate, and so the design is expected to show further compactness once we insert a dielectric substrate [13].

2) *Bandwidth:* The bandwidth (BW) of the proposed single and double band structures is reported in Tables I and II – the obtained fractional BWs (FBWs) is comparable with a 1%–4% FBW obtained from conventional single-band patch antennas [13] [39, Fig. 8]. Furthermore, we have systematically studied the tuning of the obtainable FBW. In particular, antennas with different BW can be obtained by changing the user-specified  $S_{11}$  spectrum. In the case of single-band antennas, we can sweep FBW from 0.8% to 3.3%, while in the case of double-band antennas, we can sweep FBW from 0.8% to 2.41%; Tables I and II report the highest FBW attained.

3) *Substrate Thickness:* Air-filled CP antennas were considered in [39], which reported a single-band antenna at 2.72 GHz with an FBW of  $\approx 3.33\%$  on a substrate height of 4.11 mm. On the other hand, we consider one of the proposed single-band antenna at 15 GHz (Row 3 of Table I) with the same FBW. Scaling all dimensions to match resonant frequencies, implies a substrate thickness of 3.37 mm, which is 18% thinner

than the conventional value. This implication, along with the earlier observation of a smaller area (refer to Table I) strongly favors the proposed antennas because when the structures are replaced by dielectric, the substrate losses will be lower in our structures due to the smaller antenna dimensions.

4) *Ease of Design*: Finally, we reiterate that our inverse-designed antennas have been designed in a fully automated manner near instantaneously, searching a large design space of nearly arbitrary radiating structures, whereas the other designs have either template-based geometries and therefore limited design space and functionalities, or required extensive domain knowledge (such as knowledge of slot dimensions and placement [14], [34], [35], [36]), and resource intensive optimization.

#### IV. DESIGN CONSIDERATIONS AND TRADE-OFFS

Having presented the development of the tandem network and some of the antennas generated by using our approach, we now discuss some limitations of our approach as well as possible extensions and make a prescription about antenna design in general.

##### A. Role of the Dataset Curation

As we see in Fig. 3(d), the dataset displays an imbalance at the location of the resonances. Sure enough, this imbalance seems to get reflected in the range of antennas that are generated by the tandem network as seen in Fig. 9; in particular the tandem network does not succeed in generating single band antennas in the lower end of frequencies. One of the possible ways around this is to start with a larger dataset and *filter* it in such a way as to even the imbalance; however, this is wasteful in terms of computational resources. Alternatively, including mother patches of different sizes in the dataset is also a possible way of dataset augmentation.

##### B. Role of the Substrate and Reuse of the Dataset

We have earlier referred to the design choice in the case of dataset generation—that of choosing air as the substrate for reasons of computational expediency. While this is convincing in demonstrating a methodology of design, one needs a dielectric substrate in order to realize an actual antenna. There are various possible solutions, listed below.

- 1) A brute force strategy would be to simply simulate dielectric substrates using more high-performance computing resources for dataset generation.
- 2) Another, more approximate strategy is to scale the dimensions of air-filled antennas by using an *effective* medium approach [13]. While this may work for single-band structures, the applicability in more complicated devices is not clear.
- 3) Taking a neural network approach, one can conceive of generating a smaller dataset consisting of dielectric-filled antennas and then training a network to learn the relation between the spectra of air and dielectric-filled structures. Such a network can then be cascaded to the end of the existing forward surrogate to create an end-to-end

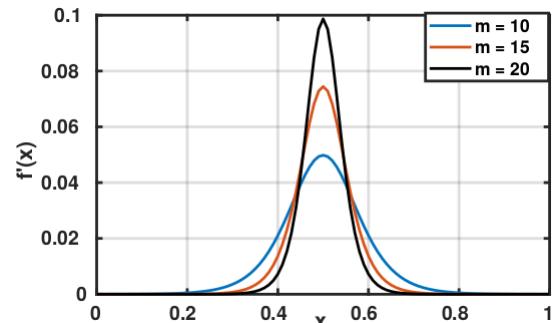


Fig. 11. Graph of the gradient of the ST activation function for different values of the ST hyperparameter,  $m$ .

network for predicting the response of a dielectric-filled structure.

- 4) Finally, the powerful idea of transfer learning [40] can be used to accelerate the training of networks that learn the dielectric response. In fact, in our parallel work [25], we show how the air-substrate dataset can be re-used for generating antennas on a Rogers RO4003C dielectric using the transfer learning approach (also fabricated and experimentally measured). In this approach, the neural network weights obtained at the end of training the air-substrate dataset are used to initialize a network that computes the dielectric-substrate response (as opposed to initializing the weights at random). The number of dielectric simulations required is drastically reduced by up to 85% (compared to the case when one would train a network from scratch) in this approach. Further, we also show how the same air-substrate dataset can be used to rapidly train networks for learning relations of different probe locations and frequency ranges.

##### C. Vanishing Gradients Problem

This article has introduced a novel ST activation function to promote discrete designs. The core reason for its success—a rapid change from 0 to 1—is also the cause of a new problem it gives rise to, that of vanishing gradients [41]. As can be seen in Fig. 11, the gradient of the ST function is nonzero only over a small range of input parameters. During network training this stalls the update of the associated parameters since the back-propagation algorithm relies on the gradient to proceed to the next step. We are thus presented with a trade-off; as seen in Fig. 7, a higher value of the ST hyperparameter  $m$  better favors discrete design, but due to the vanishing gradients problem it also leads to an increase in the training time or a failure to converge in some cases. In fact, lower values of  $m$  can lead to an overall low loss function value, but may end up generating nondiscrete and therefore *nonphysical* antennas. Several approaches have been proposed to overcome this limitation widely seen in machine learning, and remains an area of active research [42], [43].

We now discuss a couple of futuristic extensions of this work which can be expected to play a crucial role in the newly emerging area of the research described in this article.

##### D. Multiparameter Optimized Antennas

While in this article we have exclusively focused on antenna return loss as the performance parameter, the approach is

general enough to also incorporate other parameters of optimization, such as the radiation pattern, polarization ratio, gain, etc., i.e., incorporating multiple functionalities in the same structure; recent work involving templated geometries has already demonstrated this [10]. Increasing the number of parameters will come at the cost of a rise in neural network training time, which in turn forces innovation in the type of neural networks to be considered. We discuss this next.

#### E. Smarter Surrogate Models

In this (and most related work in inverse antenna or photonic structure design), the forward surrogate model is created by a data-intensive approach that involves: 1) generating data using an accurate solver and 2) learning a relationship between input and output using machine learning. Instead of a blackbox approach taken here, an emerging paradigm of “physics informed machine learning” seeks to learn input-output relations from the data, while having some of the physics of the problem (such as conservation laws or invariant relationships) be incorporated in the network architecture itself [44], [45]. As a result, these relationships need not be learned from the data, thereby reducing the amount of data required for training.

We conclude this section with a general prescription for microwave device design. While the details of this article are centered on antenna design, the approach is very general purpose and can be applied to any microwave device design in general, for e.g., power dividers/combiners, impedance transformers, filters, etc. All that is required is a suitable parameterization of the problem, followed by the standard route of data generation, and the training of the forward and tandem networks. While there is an initial cost of training, the payoff is a vast reduction in the need of domain knowledge and the near-instant availability of a design given a trained network. Such approaches are thus at a significant advantage as compared to approaches using conventional optimization-based or evolutionary algorithms which are still time-consuming as compared to all neural network approaches with run times on the order of minutes or hours [46] as compared to near instantaneous results in our approach.

## V. CONCLUSION

In this article, we have introduced a tandem neural network-based approach to design custom microstrip antennas given a desired return loss performance rapidly (in less than a second) without the requirement of extensive domain knowledge. We build on and improve existing tandem neural network approaches to allow an end-to-end deep learning-based design methodology. We highlight the innovations required to generate discrete designs by introducing a new smooth thresholding (ST) activation function and appropriate regularization terms in the loss function of the network, backed up by extensive numerical studies. The obtained antennas are significantly more compact in area as compared to conventional single- and dual-band antennas as shown with respect to existing structures in the literature. We elaborate on the various design choices encountered, highlight the shortcomings as well as the possible extensions of our work. The approach for such tandem

networks can allow rapid and effective exploration of a large design space of multifunctional antennas for future wireless systems.

## ACKNOWLEDGMENT

We acknowledge the use of the computing resources at HPCE, IIT Madras, and also the Princeton Research Computing resources at Princeton University.

## REFERENCES

- [1] H. J. Delgado and M. H. Thrusby, “A novel neural network combined with FDTD for the synthesis of a printed dipole antenna,” *IEEE Trans. Antennas Propag.*, vol. 53, no. 7, pp. 2231–2236, Jul. 2005.
- [2] Y. Kim, S. Keely, J. Ghosh, and H. Ling, “Application of artificial neural networks to broadband antenna design based on a parametric frequency model,” *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 669–674, Mar. 2007.
- [3] K. Choi, D. Jang, S. Kang, J. Lee, T. Chung, and H. Kim, “Hybrid algorithm combining genetic algorithm with evolution strategy for antenna design,” *IEEE Trans. Magn.*, vol. 52, no. 3, pp. 1–4, Mar. 2016.
- [4] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, “Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer,” *Nature Photon.*, vol. 9, no. 6, pp. 374–377, Jun. 2015.
- [5] S. Molesky, Z. Lin, A. Y. Piggott, W. L. Jin, J. Vuckovic, and A. W. Rodriguez, “Inverse design in nanophotonics,” *Nature Photon.*, vol. 12, no. 11, pp. 659–670, 2018.
- [6] J. Peurifoy et al., “Nanophotonic particle simulation and inverse design using artificial neural networks,” *Sci. Adv.*, vol. 4, no. 6, Jun. 2018.
- [7] M. H. Tahersima et al., “Deep neural network inverse design of integrated photonic power splitters,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–9, Feb. 2019.
- [8] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, “Deep learning for the design of photonic structures,” *Nature Photon.*, vol. 15, no. 2, pp. 77–90, Feb. 2021.
- [9] Y. Sharma, H. H. Zhang, and H. Xin, “Machine learning techniques for optimizing design of double T-shaped monopole antenna,” *IEEE Trans. Antennas Propag.*, vol. 68, no. 7, pp. 5658–5663, Jul. 2020.
- [10] L. Xiao, W. Shao, F. Jin, B. 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.
- [11] H. M. E. Misilmani and T. Naous, “Machine learning in antenna design: An overview on machine learning concept and algorithms,” in *Proc. Int. Conf. High Perform. Comput. Simul. (HPCS)*, Jul. 2019, pp. 600–607.
- [12] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, “Physics-informed neural networks with hard constraints for inverse design,” *SIAM J. Sci. Comput.*, vol. 43, no. 6, pp. B1105–B1132, Jan. 2021.
- [13] C. Balanis, *Antenna Theory: Analysis and Design*. Berlin, Germany: Wiley, 2015.
- [14] Z. Akhter, R. M. Bilal, and A. Shamim, “A dual mode, thin and wideband MIMO antenna system for seamless integration on UAV,” *IEEE Open J. Antennas Propag.*, vol. 2, pp. 991–1000, 2021.
- [15] D. Liu, Y. Tan, E. Khoram, and Z. Yu, “Training deep neural networks for the inverse design of nanophotonic structures,” *ACS Photon.*, vol. 5, no. 4, pp. 1365–1369, Apr. 2018.
- [16] M. Lamsalli, A. El Hamichi, M. Boussouis, N. A. Touhami, and T. Elhamadi, “Genetic algorithm optimization for microstrip patch antenna miniaturization,” *Prog. Electromagn. Res. Lett.*, vol. 60, pp. 113–120, 2016.
- [17] L. Xiao, W. Shao, F. Jin, and B. Wang, “Multiparameter modeling with ANN for antenna design,” *IEEE Trans. Antennas Propag.*, vol. 66, no. 7, pp. 3718–3723, Jul. 2018.
- [18] E. Zhu, Z. Wei, X. Xu, and W. 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.
- [19] E. Zhu, E. Li, Z. Wei, and W. Yin, “Adversarial-network regularized inverse design of frequency-selective surface with frequency-temporal deep learning,” *IEEE Trans. Antennas Propag.*, vol. 70, no. 10, pp. 9460–9469, Oct. 2022.

- [20] G. Qi, S. Bao, Y. Wei, and Y. Zhang, "Accurate antenna design by deep auto-encoder surrogate model assisted particle swarm optimization," in *Proc. IEEE 5th Int. Conf. Electron. Inf. Commun. Technol. (ICEICT)*, Aug. 2022, pp. 875–880.
- [21] D. Vercruyse, N. V. Sapra, L. Su, R. Trivedi, and J. Vučković, "Analytical level set fabrication constraints for inverse design," *Sci. Rep.*, vol. 9, no. 1, pp. 1–7, Jun. 2019.
- [22] S. K. Goudos, C. Kalialakis, and R. Mittra, "Evolutionary algorithms applied to antennas and propagation: A review of state of the art," *Int. J. Antennas Propag.*, vol. 2016, pp. 1–12, 2016.
- [23] B. Liu, M. O. Akinsolu, N. Ali, and R. Abd-Alhameed, "Efficient global optimisation of microwave antennas based on a parallel surrogate model-assisted evolutionary algorithm," *IET Microw., Antennas Propag.*, vol. 13, no. 2, pp. 149–155, Feb. 2019.
- [24] J. P. Jacobs, "Accurate modeling by convolutional neural-network regression of resonant frequencies of dual-band pixelated microstrip antenna," *IEEE Antennas Wireless Propag. Lett.*, vol. 20, no. 12, pp. 2417–2421, Dec. 2021.
- [25] E. A. Karahan, A. Gupta, U. K. Khankhoje, and K. Sengupta, "Deep learning based modeling and inverse design for arbitrary planar antenna structures at RF and millimeter-wave," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Jul. 2022, pp. 499–500. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9887077>
- [26] X. Xu, C. Sun, Y. Li, J. Zhao, J. Han, and W. Huang, "An improved tandem neural network for the inverse design of nanophotonics devices," *Opt. Commun.*, vol. 481, Feb. 2021, Art. no. 126513.
- [27] S. An et al., "Wideband Schiffman phase shifters designed with deep neural networks," in *Proc. Photon. Electromagn. Res. Symp. (PIERS)*, Apr. 2022, pp. 78–80.
- [28] C. Qiu et al., "Simultaneous inverse design continuous and discrete parameters of nanophotonic structures via back-propagation inverse neural network," *Opt. Commun.*, vol. 483, Mar. 2021, Art. no. 126641.
- [29] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2011, pp. 52–59.
- [30] L. Zhang, "Implementation of fixed-point neuron models with threshold, ramp and sigmoid activation functions," in *Proc. IOP Conf. Mater. Sci. Eng.*, vol. 224, no. 1. Bristol, U.K.: IOP Publishing, 2017, Art. no. 012054.
- [31] T. Dozat, "Incorporating Nesterov momentum into Adam," in *Proc. 4th Int. Conf. Learn. Represent.*, 2016, pp. 1–4.
- [32] L. Liu et al., "On the variance of the adaptive learning rate and beyond," 2019, *arXiv:1908.03265*.
- [33] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.*, vol. 18, no. 10, pp. 6570–6576, Oct. 2018.
- [34] A. G. Ambekar and A. A. Deshmukh, "E-shape microstrip antenna for dual frequency WLAN application," *Prog. Electromagn. Res. C*, vol. 104, pp. 13–24, 2020.
- [35] J. Chen, J. Wang, K.-F. Tong, and A. Al-Armaghany, "A GPS/Wi-Fi dual-band arc-shaped slot patch antenna for UAV application," in *Proc. Loughborough Antennas Propag. Conf. (LAPC)*, 2013, pp. 490–493.
- [36] H.-C. Tung and K.-L. Wong, "A shorted microstrip antenna for 2.4/5.2 GHz dual-band operation," *Microw. Opt. Technol. Lett.*, vol. 30, no. 6, pp. 401–402, 2001.
- [37] A. H. El-Maleh, A. T. Sheikh, and S. M. Sait, "Binary particle swarm optimization (BPSO) based state assignment for area minimization of sequential circuits," *Appl. Soft Comput.*, vol. 13, no. 12, pp. 4832–4840, Dec. 2013.
- [38] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," *Mach. Learn.*, vol. 3, nos. 2–3, pp. 95–99, 1988.
- [39] J. Eichler, P. Hazdra, M. Capek, and M. Mazanek, "Modal resonant frequencies and radiation quality factors of microstrip antennas," *Int. J. Antennas Propag.*, vol. 2012, pp. 1–9, 2012.
- [40] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, pp. 1–40, 2016.
- [41] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 1310–1318.
- [42] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [43] M. Roodzchilid, J. Gotay Sardiñas, and A. Will, "A new approach for the vanishing gradient problem on sigmoid activation," *Prog. Artif. Intell.*, vol. 9, no. 4, pp. 351–360, Dec. 2020.
- [44] W. Chen, Q. Wang, J. S. Hesthaven, and C. Zhang, "Physics-informed machine learning for reduced-order modeling of nonlinear problems," *J. Comput. Phys.*, vol. 446, Dec. 2021, Art. no. 110666.
- [45] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Rev. Phys.*, vol. 3, no. 6, pp. 422–440, 2021.
- [46] S. Song and R. D. Murch, "An efficient approach for optimizing frequency reconfigurable pixel antennas using genetic algorithms," *IEEE Trans. Antennas Propag.*, vol. 62, no. 2, pp. 609–620, Feb. 2014.

**Aggraj Gupta** (Student Member, IEEE) is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India.

**Emir Ali Karahan** (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA.

**Chandan Bhat** is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India.

**Kaushik Sengupta** (Senior Member, IEEE) received the B.Tech. and M.Tech. degrees in electronics and electrical communication engineering from the Indian Institute of Technology Kharagpur, Kharagpur, India, in 2007, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 2008 and 2012, respectively.

He joined the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, as a Faculty Member, in 2013, where he is currently an Associate Professor. His current research interests include high-frequency ICs, electromagnetics, and optics for various applications in sensing, imaging, and high-speed communication.

**Uday K. Khankhoje** (Senior Member, IEEE) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology Bombay, Mumbai, India, in 2005, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 2010.

He is currently an Associate Professor of electrical engineering with the Indian Institute of Technology Madras, Chennai, India, where he leads the numerical electromagnetics and optics (NEMO) research group which focuses on solving electromagnetics inspired inverse problems.