This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

Transfer Learning Based Rapid Design of Frequency and Dielectric Agile Antennas

Aggraj Gupta, *Graduate Student Member, IEEE* Uday K Khankhoje, *Senior Member, IEEE*

***Abstract*—Deep learning frameworks are gaining prominence in the electromagnetics community for designing microwave and mm-wave devices. This paper presents a computationally efficient transfer learning technique for designing and scaling multi-band microstrip antennas to a desired dielectric and frequency of interest. The proposed methodology involves a two-step process. First, a pre-trained model trained extensively on air-filled microstrip antennas is used for knowledge transfer. This pre-trained model is fine-tuned with a limited set of dielectric simulations, reducing data acquisition costs. In the second step, the developed forward model serves as a surrogate to design dielectric-filled antennas using the Improved Binary Particle Swarm Optimization algorithm. In contrast to conventional methods, this approach enables the design of compact antennas across various dielectrics and frequency ranges, with a significantly reduced number of time-consuming dielectric simulations (88% fewer simulations) and a lower neural network training time (75% lesser time). We analyze the optimal ways of generating dielectric antenna datasets via scaling, and perform sensitivity analysis with respect to the antenna's physical parameters. We report simulation and experimental results for single and double band antennas fabricated using the proposed approach.**

***Index Terms*—Antennas, Microstrip antennas, Microwave antennas, Multi-frequency antennas Design automation, Optimization methods, Artificial intelligence, electromagnetic theory**

I. Introduction

Modern radio frequency (RF) systems are getting increasingly compact and sophisticated, driven by innovations in computational modeling. This innovation is made possible by investing an upfront cost in terms of electromagnetic simulations that goes towards building surrogate models. These models can reasonably approximate the performance of a suitably parameterized device in a fraction of the time that a conventional electromagnetic simulation would take. The surrogate models can then be employed by a host of other techniques for the purpose of device design. These techniques include conventional gradient-based optimization algorithms, evolutionary algorithms, and machine-learning based approaches.

Here, we focus on machine-learning based surrogate models and their use in antenna design. There are broadly two approaches used in the literature to characterize a target antenna: the design could be parameterized either by means of a template, or a general pixelated geometry. The surrogate is required to learn its response, which can be in terms of the antenna return loss (S_{11}) as a function of frequency, the radiation pattern at one or more frequencies, etc. Template based approaches [1]–[5] feature a smaller input dimension to the training neural network, and hence are very economical to train in terms of the size of the training dataset and corresponding training time. However, the devices generated by them are limited by the base template, i.e. it will typically not be possible to capture all possible designs with a given template. Further, designing an appropriate template also requires a degree of specialized knowledge. On the other hand, pixelated designs [6]–[9] offer the promise of being able to express a much larger variety of antenna designs without the need of extensive domain knowledge. However, a higher up-front training and data acquisition cost is the price to be paid. This is because a larger input dimension leads to a larger number of neural-network parameters and therefore a larger training dataset size, and a correspondingly higher network training time. Such pixelated structures have also been recently explored in the design of nanophotonic devices [10]–[12].

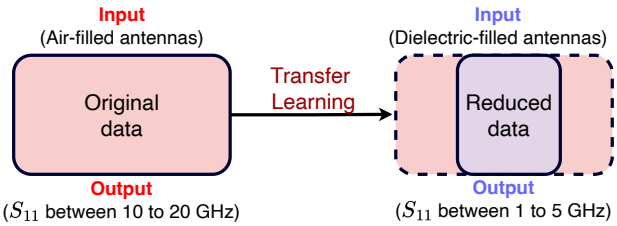

Fig. 1. An overview of the transfer-learning based framework to efficiently scale air-filled antennas across diverse dielectric and frequency ranges, leveraging a pre-trained network originally trained on air-filled samples. The proposed approach is computationally efficient and reduces the training data significantly.

In our recent work [7], we proposed a framework for antenna design that used two neural networks in a tandem configuration; the first network acted as a surrogate electromagnetic simulator, while the second network played the design role. The training dataset consisted of a large number of air-filled patch-like structures with pixelated parameterization. The choice of air-filled substrates led to a significant reduction in the training dataset collection due to the use of surface integral solvers as opposed to volume integral solvers. However, this choice left open a significant issue of the practicality of device fabrication. In this work, we propose a general framework to close this gap by using ideas from transfer

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India

Corresponding author: Uday K Khankhoje (e-mail: uday@ee.iitm.ac.in).

Authors acknowledge the following research grants: (i)"6G: Sub-THz Wireless Communication with Intelligent Reflecting Surfaces (IRS)," R-23011/3/2022-CC&BT-MeitY by the Ministry of Electronics and Information Technology (MeitY), Government of India, and (ii) "Intelligent approaches to designing and building Intelligent Reflecting Surfaces for 6G applications," by the Qualcomm 6G University Research India Program.

learning [13].

A. Related work

The use of transfer learning to extend the capabilities of electromagnetic surrogates is relatively new. In a related study [14], a transfer learning approach was employed for the inverse design of functional metasurfaces, treating the phase-to-pattern inverse design as a classification problem; here the transfer learning model was based on the Inception V3 framework [15], [16]. Recent work [17] briefly touched upon transfer learning approach to scale microwave structures, specifically output matching networks for power amplifiers for different dielectric stacks; however, the study did not quantify this approach in terms of dataset generation and acquisition costs, the scaling factors involved, sensitivity analysis or the achieved data reduction through transfer learning. Qu et al. [18] demonstrated the possibility of predicting optical properties in a given physical scenario with the help of the knowledge obtained from a different, but related, physical problem. A recent study by Fan et al. [19] demonstrated the effectiveness of knowledge transfer learning in the inverse design of metasurfaces. The authors successfully migrated the inverse design process from a 5 $\times$ 5 metasurface to a 20 $\times$ 20 metasurface. More recently, Kiani et al. [16] have shown how reconfigurable electromagnetic metasurfaces can be designed and tuned to achieve multiple functionalities. In [20], the authors presented an inverse design of slow wave structures, extending from the *Ka* band to a center frequency of 850 GHz. In [21], a novel inverse design model using a Generative Adversarial Network (GAN) was developed for a varactor-based tunable lowpass filter (TLPF), leveraging transfer learning. The research in [22] introduces an efficient transfer learning method that allows for the reuse of design knowledge across different metasurface applications, reducing the data needed and broadening the possibilities for metasurface design. Lastly, the work in [23] focuses on the development of a transfer learning-based deep neural network (TL-DNN) model for the inverse, fully-automated design of on-chip interconnects.

B. Our contributions

We propose a transfer learning framework trained on Task A (surrogates for EM simulations of air-filled antennas) to learn how to do Task B (surrogates for EM simulations of dielectric-filled antennas) or learn how to do Task C (surrogates for EM simulations of air or dielectric filled structures in different frequency ranges than Task A). Our proposed framework capitalizes on a pre-trained network (trained on air-filled substrates) developed in our previous work [7] and by additionally investing in a much smaller dataset of dielectric-filled samples, we formulate a technique to design antennas in different frequency bands and dielectric substrates as per user specifications (see summary in Fig. 1). Compared to training a surrogate neural network from scratch, this approach offers significant computational gains and flexibility. To the best of our knowledge, no prior work has leveraged a generalized deep-learning-based architecture for designing antennas across any desired dielectric and frequency span.

In this study, we examine each aspect extensively through numerical studies using microstrip patch antennas as a proof of concept to validate our synthesis method. Additionally, we scale our structures (here, antennas) across both dielectric stacks and frequency spans. We note that a detailed mathematical exposition on knowledge transfer has been covered elsewhere, for e.g. [24], [25], and hence is not covered here. Instead, we utilize it as a technique for developing computationally effective surrogate models for inverse antenna design.

The paper is organized as follows. In Section II, we define the problem by describing the transfer-learning surrogate model developed by us and the approach for the inverse design of antenna structures. In Section III, we elaborate on the details of the datasets required for network training and present an analysis with of optimal network performance with regards to dataset sizes. In Section IV, we report the various types of single and double band antennas designed by our approach, and analyze the sensitivity of the inverse-designed structures. In Section V, we report experimental results of our fabricated devices, and finally conclude in Section VI.

II. Methods

In this Section, we provide an overall sketch of the proposed methodology for antenna design. Formally, the problem statement we solve in this work is as follows: Efficiently design a multi-band antenna on a given dielectric substrate and at given frequency bands. We are given access to a trained neural network which simulates the response of a pixelated air-filled substrate microstrip antenna over a pre-specified frequency range.

A. Solution strategy

As an illustration of the challenge involved, we compare the antenna return loss for an air-filled antenna, and the corresponding response when a dielectric substrate is added to the same structure. We consider such a situation in Fig. 2, which shows that the response of the air and dielectric substrate bear no relation to each other. Therefore, at face value, one might not consider the air-filled simulations to be of much relevance to the problem at hand.

However, one must note that the antenna responses shown in Fig. 2 are the responses of a full wave electromagnetic solver. In practice, using such a solver for the purpose of antenna *design* is very impractical due to the large simulation times involved. For example, if we were to use an evolutionary algorithm for antenna design, it would repeatedly invoke the electromagnetic solver to evaluate the fitness of a given design before finally converging on a design. Due to their extremely fast evaluation time, neural-network based surrogates of electromagnetic solvers are beginning to play a game changing role in design problems.

Thus, it is essential to build a neural network to simulate the response of dielectric-filled substrate. The crucial question is about which route is the most *efficient* in terms of computational resources. The naive approach of training such a network from scratch is quite expensive given the fact that the

This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

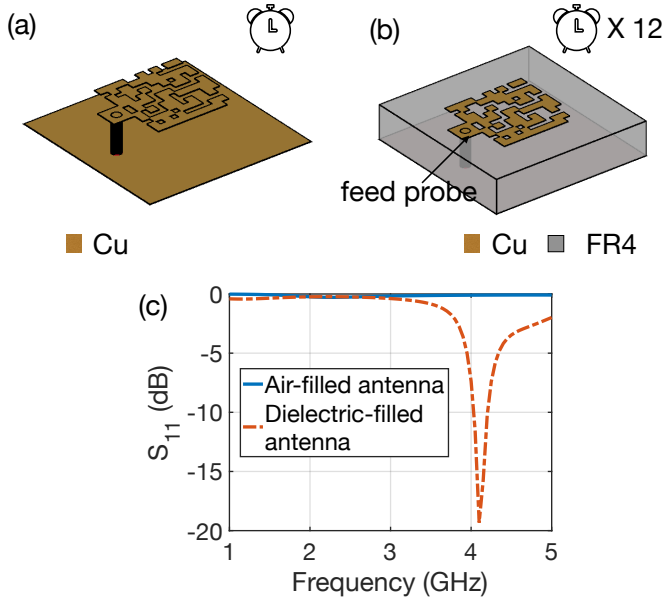

Fig. 2. (a) A sample air-filled antenna, and (b) a dielectric-filled antenna, with the corresponding return loss (S_{11}) plotted in (c) using MATLAB's Antenna toolbox.

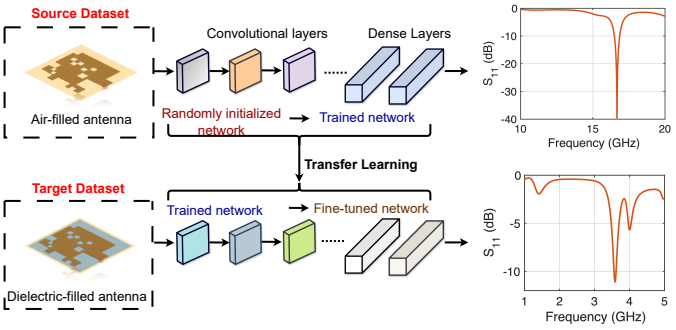

Fig. 3. Transfer learning assisted forward surrogate model

dataset generation time will be very high. This is because a dielectric-filled antenna simulation takes approximately twelve times more compute time than a corresponding air-filled antenna simulation when using the method of moments electromagnetic solver. An air-filled antenna simulation is particularly fast because the air-region of the problem does not require meshing, only the surface does.

Instead of the naive approach, we adopt the transfer learning strategy since we already have access to a pre-trained network that simulates the response of an air-filled antenna structure [7] based on a large dataset of size 500k. We take the same neural network architecture as before and instead of initializing the network weights at random (as is usually done), we load the weights from the pre-trained network as shown in Fig. 3. Then, using a small, newly generated dataset of dielectric antenna simulations, we further train this network. As we show subsequently in the results, this approach leads to much faster network convergence as compared to training from random initialization.

Having built a fast surrogate model, we proceed to the task of antenna design itself. We use the well established Improved Binary Particle Swarm Optimization [26] algorithm for this task. The algorithm starts with initializing particles and then successively updating the velocity, position, local best and global best solutions with each iteration. The fitness is evaluated by calculating the fitness for each initialization which requires invoking the newly built surrogate model. For a given structure with return loss (in dB) characterized at discrete frequency points, i.e. $\{S_{11}(f_i)\}$, the fitness function is evaluated as:

$$g = \sum_{j \in P} S_j + \sum_{j \in S} (10 - S_j), \tag{1}$$

with,

$$S_j = \begin{cases} |S_{11}(f_j)| & 5 \leq |S_{11}(f_j)| \leq 10 \\ 10 & |S_{11}(f_j)| \geq 10 \\ 0 & |S_{11}(f_j)| \leq 5 \end{cases} \tag{2}$$

where P and S are the sets of pass-band and stop-band frequencies, respectively. After reaching the maximum iteration value, the individual with the greatest fitness obtained during the evolution process is reported as the optimal design.

We note in closing that instead of an evolutionary approach, we could have just as well used an all-neural network approach, such as the tandem network from our previous work [7], to design the dielectric-filled antennas. However, we choose to limit the focus of this study to the efficient creation of surrogate models for dielectric antennas using transfer learning and measurements of antennas fabricated using them.

III. Transfer Learning Assisted Surrogate Electromagnetic Model

In this Section, we present details of the neural network used for training the surrogate model, generating the dataset, and other network training details.

A. Antenna scaling and dataset generation

We initiate the antenna design by tessellating a square metallic patch area into a 12×12 grid of pixels. A design is specified by each pixel taking on a value of 1 or 0, signifying the presence of metal or its absence, respectively. We fix the two pixels near the feed location to always be metal to ensure connectivity, thus giving us an extensive design space of 2^{142} potential structures.

The next highly critical step is to get the approximate dimensions of the "mother" patch for the corresponding choice of dielectric substrate, height, and choice of design frequency. We do this by judiciously scaling the length and width of antenna structures from the air-filled dataset to get approximately viable radiating structures for populating the "dielectric" dataset. This scaling is necessary since we can see from Fig. 2 that simply inserting a dielectric into an existing air-filled design does not lead to a radiating structure.

The air-filled structures from our training dataset in [7] had a substrate thickness of 0.61 mm and resonated in the 10-20 GHz range. On the other hand, for the purpose of this study, we choose to design and fabricate antennas in the 1-5 GHz band using a standard commercially available $h = 3.2$ mm FR4 substrate having ϵ_r and loss tangent ($\tan\delta$) as 4.8 and 0.026 respectively.

This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

B. Effect of scaling factor on the resonant samples in the dataset

Since the new substrates contain a dielectric material instead of air, we use the concept of effective relative permittivity to model the resonant frequency of the structure. This concept is commonly used in the case of microstrip patch antennas [27] to relate the patch dimensions to the effective relative permittivity as: $\epsilon_{reff} = 0.5[(\epsilon_r+1)+(\epsilon_r-1)(1+12h/w)^{-0.5}]$. The resonant frequency for the lowest TM_{010}-like cavity mode has a dependence on width, w, and effective permittivity as $f \propto (w\sqrt{\epsilon_{reff}})^{-1}$. Thus, the width of the dielectric filled antenna, w_d, can be related to that of the air-filled antenna, w_a, as:

$$w_d = \left\{ \frac{f_a}{f_d} \frac{1}{\sqrt{\epsilon_{reff}}} \right\} w_a = a\, w_a \tag{3}$$

where f_a, f_d are the resonant frequencies of air and dielectric filled antennas, respectively, and we refer to a as the scale factor. Since ϵ_{reff} itself depends on w_d, we solve the nonlinear equation suggested in Eq. (3) to obtain the scale factors to translate the air filled antennas from 10-20 GHz to the 1-5 GHz range in FR4. These scale factors are found to lie in the range, 2 to 5.

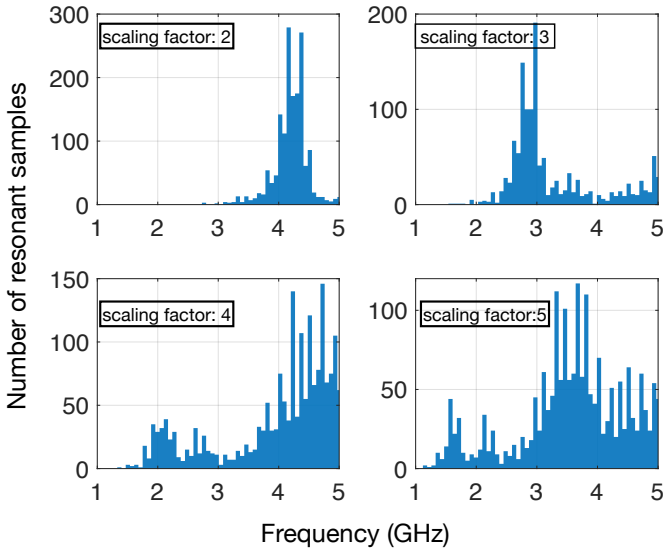

Fig. 4. Histogram plot for different scaling factors, a.

The histogram plot in Fig. 4 visually demonstrates the impact of scaling structures from the air-filled dataset. An increase in the scaling factor reduces the resonant frequencies due to an increase in patch dimensions. The figure also suggests that higher scale factors lead to a broader range of resonances. We have previously shown that having a diversity in the training dataset is critical to allow for inverse design within a desired frequency range [7]. Thus, while higher scale factors provide us with greater dataset diversity, they present a trade-off for device compactness since higher scale factors correspond to larger dimensions.

We note that the expression in Eq. (3) gives us an approximate scale factor. However, this expression was derived for conventional rectangular patch antennas [27] and will not strictly hold for pixelated patch antennas. Therefore, since these scale factors are approximate, we explore pushing these limits for our unconventional designs. A lower scale factor corresponds to a more compact design and hence we choose a scaling factor of 1.94 in this work. This choice highlights the compactness of our inverse-designed structures in the 5G communication n78 frequency band, covering 3.4 to 3.8 GHz.

We employed MATLAB's Method of Moment (MoM) based solver with a mesh size of $\frac{\lambda}{15}$, where λ represents the guided wavelength at the design frequency. The return loss of dielectric-filled antennas was simulated for 81 equidistant points across the frequency range of 1 to 5 GHz. Each dielectric simulation took approximately 110 seconds to complete. These simulations were performed on a 2.5 GHz Dual-Core Intel Core i5 processor with 8 GB 1600 MHz DDR3 RAM. The process of generating the dataset on a 160-core server required approximately one day. A total of 60k antennas and their spectra are generated, which are distributed in a 90:5:5 ratio for training, validation, and testing, respectively.

C. Network training

We now consider a neural network with the same forward model architecture as in [7]. The network is composed of 56 layers, incorporating filters, weights, and biases [7]. The initial 16 layers consist of a combination of 2-D convolutional layers, batch normalization, and the leaky ReLU activation function layers, as depicted in Fig. 5(a). Following these, two fully connected layers are integrated, with batch normalization, leaky ReLU activation functions, and dropout (set at a value of 0.4) applied to the output of each fully connected layer. The output of the final fully connected layer is directed to an 81-dimensional output regression layer. The loss function used to train the proposed forward model is a mean squared error function of the form:

$$L_f = \frac{1}{P} \sum_{i=1}^{P} \text{MSE}(\mathcal{S}_i, \mathcal{S}'_i) \tag{4}$$

where P is the batch size, set to 256, $\mathcal{S}$ and $\mathcal{S}'$ are the true and predicted spectrum, respectively. We choose NAdam [28] optimizer to update the weights and biases of the CNN. Operating on the input of an FR4-filled antenna design, the proposed network delivers the corresponding return loss response as its output. We trained the forward network using PyTorch on the Google Colab GPU platform, carrying out 50 epochs which took approximately 45 minutes. This is in contrast to the 3 hours training time taken in our previous work [7], i.e. we achieve a 75% reduction in training time in this work. The recorded training and validation losses were 0.18 and 0.38, respectively. These values are notably lower than the losses reported in our earlier publication on air-filled antennas [7], despite the significantly larger dataset size (500k). Once trained, the proposed network is capable of predicting the return loss of dielectric-filled antennas in less than a second. Figure 5(b) provides a representation of the reconstruction results obtained from the test dataset using the trained forward model.

We conducted a thorough hyperparameter (HP) search using the "WandB" tool [29] which logs all runs with different HPs

This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

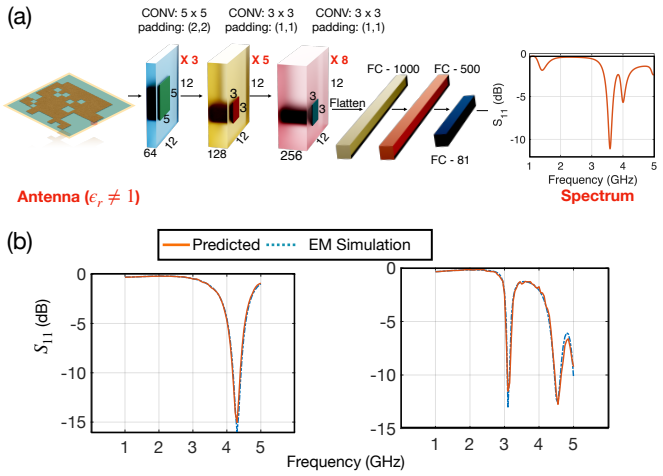

Fig. 5. (a) Transfer-learning based forward CNN consisting of convolutional layers and fully connected layers denoted by "CONV" and "FC" respectively, (b) Comparison of the predicted S_{11} spectrum from the proposed forward model with the EM simulated spectrum from electromagnetic simulator showing the efficacy of the forward model

during the modeling of deep networks. The dataset used for HP tuning (dielectric-filled structures) consists of 20k samples for training and 10k for validation. Fig. 6 shows the parallel plot for HP tuning (random search) for the proposed transfer learning surrogate model. We observe that with 50 epochs and an initial learning rate of 0.001, the validation error is minimized. Therefore, these HPs are selected. Higher learning rates, such as 0.1, result in significantly higher validation loss. Thus, for this problem, an initial learning rate of 0.001 is optimal. Additionally, it has been observed that increasing the number of epochs reduces the validation error, indicating that training the network for a sufficient duration improves performance. Another important aspect to consider is overfitting. From Fig. 6, we observe that for a minimum training loss of 0.32, overfitting occurs with batch sizes smaller than 256 (i.e., 64 and 128). Both batch sizes of 256 and 512 are good candidates for training the network. However, we choose a batch size of 256, as smaller batch sizes introduce randomness (noise) into the optimization process. This noise promotes exploration of the loss landscape, helping the model find flat minima [30]. Flat minima are preferred because they improve generalization, making the model more robust to new, unseen data.

To ensure the robustness of our surrogate model, we conducted k-fold cross-validation with $k = 5$. This helps prevent overfitting and provides a more reliable estimate of the model's accuracy across different data splits. Our results show consistent performance across all folds, with no unusual outliers that could skew the evaluation. This consistency confirms the robustness and generalizability of the model.

Table. I shows the range of values used for the random search and the chosen set of HPs to train the transfer learning based forward surrogate model. It has been observed that when the network's HPs are not chosen wisely, it results in high error rates and, consequently, inaccurate predictions. Other ways to perform HP tuning include probabilistic methods as referenced in [31]–[34].

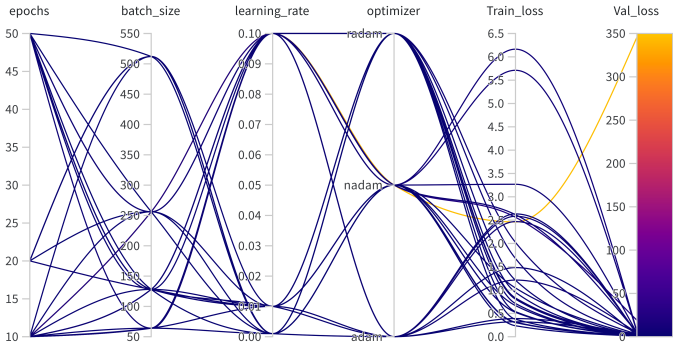

Fig. 6. Parallel plot for hyperparameter tuning generated by WandB [29], showing epochs, batch size, learning rate, and optimizer. Train and validation losses on the right are used to select optimal hyperparameters.

TABLE I
OPTIMAL SET OF HYPERPARAMETERS CHOSEN TO TRAIN THE PROPOSED TRANSFER LEARNING BASED FORWARD SURROGATE MODEL

Hyperparameter	Range of values	Optimal value
Batch size	[64, 128, 256, 512]	256
Learning rate	[0.1, 0.01, 0.001]	0.001
Optimizer	[Adam, RAdam, NAdam]	NAdam
Number of epochs	[10, 20, 50]	50

We track the validation set root mean squared error (rmse) for both cases – the first when the network is trained with random initialization (i.e. from "scratch"), and the second when we use transfer learning (TL) and initialize from the parameter values as the air-filled simulator's parameters. These network validation rmse values are shown in Fig. 7.

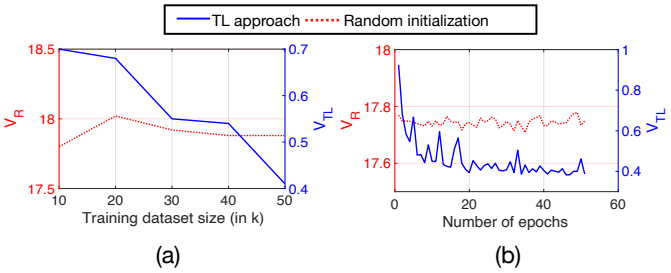

Fig. 7. Comparison of validation error convergence for training the network from scratch (V_R) and training the network using the Transfer learning (TL) approach (V_{TL}): (a) with respect to the different training set sizes, (b) with respect to the number of epochs

As we can see from Fig. 7(a), while training the networks on datasets of increasing size from 10k onwards, the validation rmse decreases steadily for the TL approach as the size of the dataset increases. The TL approach shows nearly an order of magnitude lower rmse than the "from scratch" approach; at a training dataset size of 50k, the TL approach has an rmse of 0.4, which is more than 44 times lower than the corresponding rmse for the other approach. While training the CNN with a dataset size of 60k, it becomes apparent that the validation error for the randomly initialized network remains nearly constant and does not decrease with epochs (refer to Fig. 7(b)). In contrast, the validation error using the

This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

TL approach converges to a significantly low value. Thus, our proposed methodology converges to a lower validation loss with significantly lesser data, in contrast to the randomly initialized network. In our work, we use a dataset size of 60k for antenna design, which is significantly smaller than the air-filled dataset (88% lesser). That said, the two datasets have comparable generation times. The key observation is that it is *only* due to the TL approach that the new dataset suffices for antenna design since it is leveraging the learning from the original dataset.

IV. Antenna Design: Simulation Results

In this section, we report simulation studies of various antenna structures scaled on FR4 substrates in the 5G communication n78 frequency band using our proposed approach.

A. Single band antennas

Fig. 8 illustrates a compact inverse-designed antenna, operating within the frequency range of 3.55 - 3.65 GHz, with the center frequency 3.6 GHz. As seen in Fig. 8(b), the antenna spectrum is in good agreement with the design specification and the corresponding full wave electromagnetic simulation. The entire design process for this structure required approximately 15 minutes for the Improved BPSO algorithm.

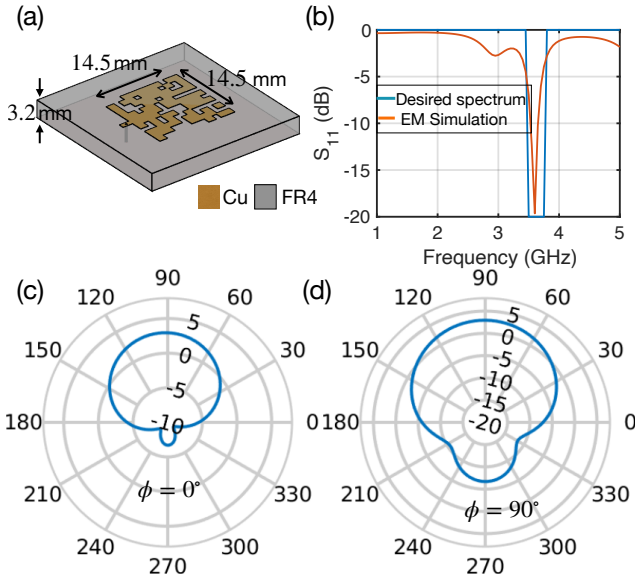

Fig. 8. (a) Inverse designed single-band antenna, (b) reconstructed spectrum, (c) Elevation radiation pattern ($\phi = 0°$) at 3.6 GHz, (d) Elevation radiation pattern ($\phi = 90°$) at 3.6 GHz

The axial ratio of the reported antenna is $\geq$ 25 dB for $\theta = \pm 90°$ for both the planes ($\phi = 0°$ and $\phi = 90°$) from the main lobe. It is noteworthy that while the axial ratio and radiation pattern were not integrated into the cost function for optimization within this study, the radiation pattern of the inverse-designed antenna exhibits a well-defined shape (Fig. 8(c) & (d)), directed at $\theta = 90°$ with a gain of 3.5 dB. Moreover, the device is characterized as linearly polarized, demonstrating a radiation efficiency exceeding 70%, accounting for both copper losses and the inherent losses within the FR4 substrate as observed in conventional microstrip antennas [27]. The reason for the broadside radiation pattern of the antenna, even though it has not been specified in the optimization process, is the low $\frac{h}{\lambda}$ ratio in the patch antenna (h: height of the substrate, λ: wavelength in the dielectric). In our study, for the antenna reported in Fig. 8(a), $h = 3.2$ mm, and $\lambda = 83.3$ mm at 3.6 GHz, hence the ratio $\frac{h}{\lambda} \ll 1$. Therefore, the radiating edges of the antenna can be thought of as wire antennas carrying magnetic current (perpendicular to the plane of the antenna), contributing to the overall broadside radiation pattern.

TABLE II
Design parameters for the single-band resonant conventional patch (CP) and the inverse designed patch (IDP) antenna for different frequencies in the dataset. CPL: Conventional patch length; CPW: Conventional patch width; f_0: Resonant frequency, Δ area: Compactness relative to CP, BW: BandWidth, FBW: Fractional BandWidth

CPL (mm)	CPW (mm)	f_0 (GHz)	Area (mm^2) (CP)	Area (mm^2) (IDP)	Δ area %	Gain (dB) CP	Gain (dB) IDP	BW (FBW) MHz (%)
17.5	20.3	3.65	355.7	210	41	4.8	3.3	100 (2.7%)
16	21.8	3.75	350.5	210	40	5	3.5	100 (2.7%)
15	22	4	320	210	36	4.4	4	100 (3.3%)
13	19.5	4.5	280	210	33	5.1	4.5	100 (2.7%)

The antenna is approximately 41% smaller compared to a conventional antenna operating at the same frequency, thus showing a significant achievement in antenna miniaturization.

We compare several single band antennas with their conventional counterparts in Table. II. Conventional patch antennas cover an area of approximately $\frac{\lambda}{2} \times \frac{\lambda}{2}$ [27], while in this work the antennas occupy an area on the order of $\frac{\lambda}{4} \times \frac{\lambda}{4}$, which is significantly compact at the design frequency. We have noticed a trade-off in compactness (expressed as a percentage) and the difference in the effective aperture area between conventional and inversely designed structures within the frequency range of interest. Attempting inverse design at the lower end of the spectrum results in the most compact devices, but sacrifices antenna gain. This is because for a given frequency, gain varies proportionally with the aperture area [27]. As we move to lower frequencies, our inverse designed devices tend to become more compact, reducing the effective aperture area and, consequently, the gain of the device. Our observations indicate that the gap in effective aperture area widens between inversely designed structures and their conventional counterparts at the lower end of the spectrum, for instance, when we attempted the inverse design for an antenna operating at 3.5 GHz (50% more compact), we were able to successfully design an antenna with a gain of 1.73 dB. In comparison, the conventional antenna at the same frequency had a gain of 4.51 dB. Therefore, to maintain comparable gain with their conventional counterparts (refer to Table II), we design antennas that are 41% more compact at the desired frequencies.

B. Dual band antennas

Fig. 9(a) shows a dual-band inverse designed structure resonating at center frequencies 3.65 GHz and 4.65 GHz, respectively. It took 40 minutes for the Improved BPSO algorithm to converge to the dual-band structure shown in Fig. 9(a). The spectrum of the inverse-designed structure closely matches the that from a full wave electromagnetic simulation of the device, as seen in Fig. 9(b).

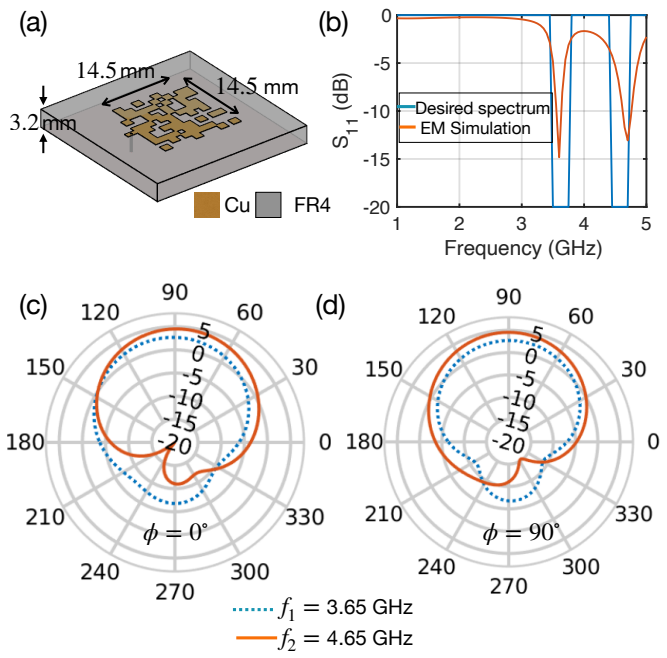

Fig. 9. (a) Inverse designed dual-band antenna, (b) reconstructed spectrum, (c) Elevation radiation pattern ($\phi = 0°$) at f_1 = 3.65 GHz and f_2 = 4.65 GHz (bd Elevation radiation pattern ($\phi = 90°$) at f_1 = 3.65 GHz and f_2 = 4.65 GHz

TABLE III
DUAL-BAND SPECTRA OF INVERSE DESIGNED ANTENNAS; HERE λ_0 DENOTES THE WAVELENGTH CORRESPONDING TO THE FIRST FREQUENCY BAND, f_1, BW: BANDWIDTH , FBW: FRACTIONAL BANDWIDTH, RE: RADIATION EFFICIENCY

f_1 (GHz)	f_2 (GHz)	Gain (dB)		BW (FBW) MHz (%)		Aperture area
		f_1	f_2	f_1	f_2	
3.6	4.5	3	2.5	80 (1.5%)	100 (1.7%)	0.030 λ_0^2
3.7	4.9	3.8	3	75 (1.4%)	90 (1.6%)	0.032 λ_0^2
3.8	4.8	4	3.34	80 (1.5%)	75(1.4%)	0.033 λ_0^2
4	4.5	4.2	3.8	250 (1.8%)	250 (1.4%)	0.037 λ_0^2

The radiation pattern for the inverse-designed antenna in Fig. 9(a) at both the design frequencies are shown in Fig. 9(c) & (d), with gain of 4 dB and 4.5 dB at 3.65 GHz and 4.65 GHz, respectively. Table. III shows different dual band designs, where we also report the area of obtained antennas in terms of the wavelength of the lower frequency band, λ_0. The axial ratio of the reported antenna is $\geq$ 20 dB for $\theta = \pm 90°$ for both the planes ($\phi = 0°$ and $\phi = 90°$) at both the frequencies from the main lobe.

TABLE IV
STATE-OF-THE-ART COMPARISON FOR DUAL-BAND MICROSTRIP ANTENNAS. NR: NOT REPORTED

Reference	f_1 (GHz)	f_2 (GHz)	Gain (dB)		BW (MHz)		Aperture area
			f_1	f_2	f_1	f_2	
[35]	4.53	4.97	5	4.97	NR	NR	1.72 λ_0^2
[36]	2.4	5.6	9.55	8.50	97	360	1.16 λ_0^2
[37]	2.4	5.8	6.8	2.1	164	256	0.64 λ_0^2
[38]	1.57	2.45	7.47	7.07	47	125	0.64 λ_0^2
[39]	2.4	5.2	4.1	1.4	90	277	0.08 λ_0^2
[40]	3.45	5.9	3.83	0.57	140	150	0.17 λ_0^2
This work	3.65	4.65	3.7	4.3	100	125	0.03 λ_0^2

We have compared and reported the average run-time (over ten runs each) of the deep learning-based surrogate model assisted evolutionary algorithm to get the desired response, as shown in Fig. 8 (single-band) and Fig. 9 (dual-band). Table. V (similar to Table. I in [41]) shows the performance of the proposed approach as compared to our previous work [7]. We have used the same neural-network forward model architecture as a surrogate for inverse design of the antennas for a fair comparison.

C. Sensitivity analysis

To study the impact on the performance of our devices due to the inherent imprecision in the fabrication processes and variations in the permittivity and thickness of the FR4 substrate, we perform a sensitivity analysis. Specifically, we study $\pm 1\%$ variations in the size of the patch, $\pm 10\%$ variations in substrate thickness, and $\pm 10\%$ changes in the relative permittivity (ϵ_r). For definiteness, we consider the single-band antenna shown earlier in this Section, resonating at 3.65 GHz with a bandwidth of 100 MHz. We compute sensitivity as $\frac{\|S-S'\|}{\|S\|} \times 100$, where S represents the true response of the

TABLE V
PERFORMANCE COMPARISON OF THE PROPOSED APPROACH FROM OUR PREVIOUS WORK [7] IN TERMS OF THE COMPUTATIONAL COST, DATASET REQUIREMENT, HYPERPARAMETER TUNING AND OTHER VARIOUS PARAMETERS. FM: FORWARD MODEL, IDM: INVERSE DESIGN METHOD, CT: COMPUTATIONAL TIME, TL" TRANSFER LEARNING, NN: NEURAL NETWORK, HP: HYPERPARAMETER, SB: SINGLE BAND, DB: DUAL BAND

Parameters	Previous work [7]	Proposed work (TL)
FM	NN	NN
IDM	NN	Improved BPSO
Antenna type	Air filled	Dielectric filled
CT for training (one time)	3 hrs (FM) + 12 hrs (IDM)	45 min (FM)
Dataset requirement	500k	60k
Antenna Design time	$\leq$ 1 sec (SB & DB)	SB (15 mins), DB (40 mins)
Number of solutions per input	Single solution	Different solution per run (Explores search space better)
HP for inverse design	Extensive HP tuning required	No HP tuning is required

This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

structure, and S' corresponds to the response while varying one parameter (patch size, permittivity, or thickness) while keeping the other two constants. Monte Carlo simulations were employed for sensitivity analysis on all parameters, randomly selecting values within specified variations and performing full-wave simulations; convergence was observed in 200 simulations for each parameter.

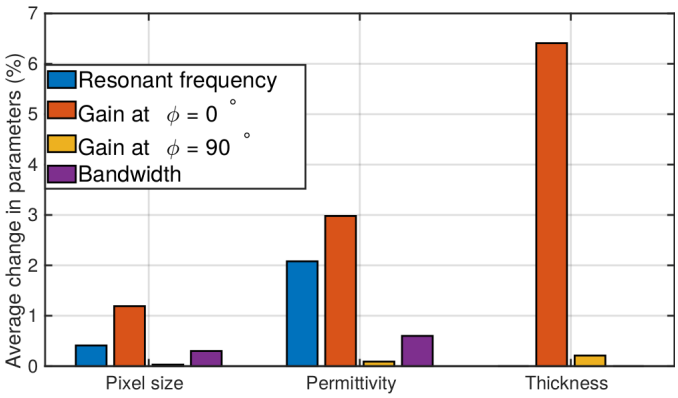

Fig. 10. Sensitivity Analysis plot for the single-band inverse designed antenna at 3.65 GHz. The plot shows the sensitivity of the variation of resonant frequency and gain with respect to change in substrate thickness ($\pm 10\%$), permittivity ($\pm 10\%$), and patch size ($\pm 1\%$).

The results of the analysis are shown in Fig. 10. As can be seen, the parameter most susceptible to changes is the gain at ϕ = 0° ($\pm 6\%$), particularly due to the change in the permittivity ($\pm 10\%$). On the other hand, variations of $\pm 10\%$ in substrate thickness do not impact the resonant frequency. However, these changes do affect the gain. The analysis [27, Ch. 14.27] reveals that gain tends to increase initially with an increase in substrate thickness up to a certain threshold, primarily due to a decrease in the capacitance value. However, beyond the certain threshold of substrate thickness, the radiation efficiency begins to decrease, mainly because of the emergence of surface waves within the dielectric material [42]. We conclude from Fig. 10 that our proposed inverse design structures are robust to the changes in substrate thickness and permittivity as well as the pixel size.

D. Comparison with conventional antennas

Aperture area: We refer the reader to Table IV, where a comparison of various dual band structures is presented. In particular, our designs yield much more compact antennas than those reported in the literature. For instance, the lowest area achieved by our work is $0.03\lambda_0^2$, while these are the areas reported in other works: $1.16\lambda_0^2$ for a dual mode circular patch antenna in [36], $0.64\lambda_0^2$ for an E-shaped microstrip patch in [37] and an arc-shaped slot patch antenna in [38], and $0.08\lambda_0^2$ for a shorted microstrip antenna in [39]. It must be noted in the latter case that the lower aperture area is coming at the cost of significantly lower antenna gain than our antennas.

Bandwidth: The bandwidth of the proposed single and double band structures is reported in Tables II and III – the obtained fractional bandwidths (FBW) is comparable with a 1-4% FBW obtained from conventional single-band patch antennas [27] [43, Fig. 8]. Since the current work is centered on designing antennas within the 5G communication n78 frequency band (3.4 - 3.8 GHz), it is crucial to acknowledge that, in this densely populated frequency range, the maximum channel bandwidth specified for this band of spectrum is 100 MHz. Our inversely designed structures can effectively operate within the entire bandwidth allocated to the n78 frequency band (refer to Table II).

Polarization stability: Similar to traditional antennas [27], our inverse-designed structures exhibit linear polarization. Specifically, considering one of the proposed single-band antennas at 4 GHz (Row 3 of Table II), the axial ratio is approximately 26 dB at the broadside direction for both the planes. Further, by randomly inspecting 1000 samples from the dataset, the axial ratio is found to be a minimum of 28 dB, greater than the 20 dB threshold for the structure to be linearly polarized [44]. Therefore, the polarization of the pixelated antennas remains linear even though the geometry changes randomly.

V. Antenna Design: Experimental Results

To validate the inverse-design assisted by the transfer-learning based approach, we fabricated several antennas across the 3-5 GHz frequency range in a 3.2 mm FR4 substrate with 3.2 mm. To connect the antennas to the dielectric's edge using a microstrip feed line, we extend the inverse-designed antennas through a microstrip feed line with a width of 6.27 mm. This specific width is chosen to achieve a 50-ohm impedance for the microstrip line, considering the selected thickness and dielectric material.

The S-Parameter measurements were performed with Keysight's PNA-X Network Analyzer (N5244B). We have also measured the radiation pattern of the fabricated antennas in transmit mode in an anaechoic chamber as per the setup shown in Fig. 11.

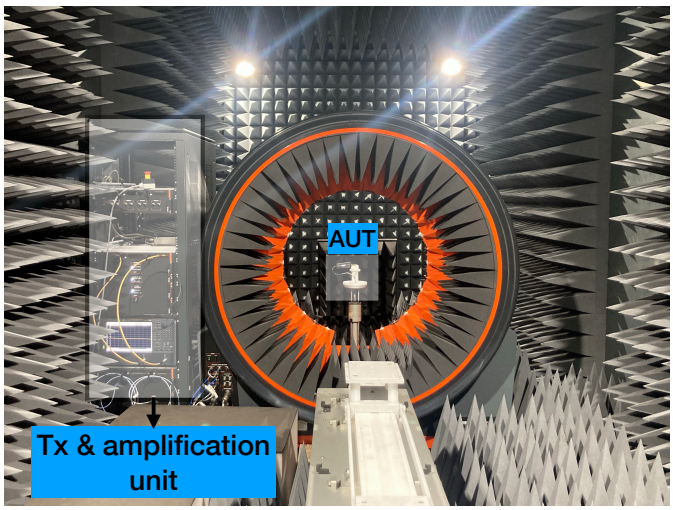

Fig. 11. Radiation pattern setup for the measurement of gain for the inverse-designed structures

Fig. 12(a) shows a fabricated edge-fed 3.75 GHz single band antenna synthesized by our approach. The measured

This article has been accepted for publication in IEEE Journal on Multiscale and Multiphysics Computational Techniques. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JMMCT.2024.3509773

S-parameters match well with the simulated results, as can be see in Fig. 12(b). The antenna dominantly radiates in

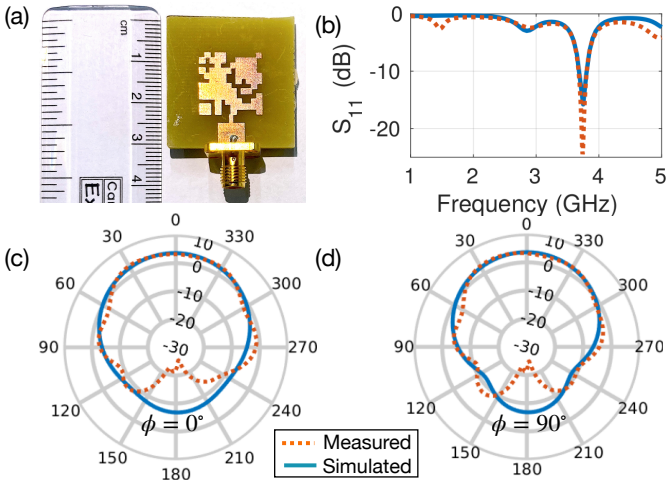

Fig. 12. (a) An instance of fabricated inverse designed single-band antenna, (b) reconstructed spectrum vs the measured spectrum at 3.75 GHz, (c) Elevation radiation pattern ($\phi = 0°$) at f_1 = 3.75 GHz (d) Elevation radiation pattern ($\phi = 90°$) at f_1 = 3.75 GHz

the broad-side direction at the frequency of interest, and the measured patterns show good unidirectional radiation profile and symmetry in the $\phi = 0°$ and $\phi = 90°$ planes, as seen in Fig. 12(c,d). Additionally, the patterns match fairly well with the simulation results. We observe a gain of approximately 3.2 dB at the resonant frequency with a fractional bandwidth of 4%, which is comparable to conventional microstrip antennas [27].

Additionally, alongside presenting the radiation pattern, we include co-polarisation and cross-polarization measurements (Fig. 13) for the fabricated inverse design antenna depicted in Fig. 12(a). We note that the measured cross-polarization isolation stands at around 7-8 dB, and ideally this isolation should be higher. This is attributed to using a square mother patch in this study and a substrate thickness of 3.2 mm. It is known [45]–[47] that cross-polarization isolation increases when the length-to-width ratio of the patch antenna is approximately 1.6, favouring rectangular patches over square ones. Furthermore, thinner substrates exhibit better isolation than thicker ones [45], given the optimal length-to-width ratio. Therefore, this limitation observed in the inverse-designed structure is not fundamental and can be easily further improved by carefully selecting the mother patch's optimal length, width, and thickness.

Fig. 14(a) shows an example of a fabricated edge-fed dual band antenna resonating at 4 and 4.7 GHz. The measured S-parameters in Fig. 14(b) show the dual resonances as expected, and the measured radiation patterns at both frequencies demonstrate the desired unidirectional patterns with 4.12 dB and 4.31 dB gain at frequencies 4 GHz and 4.7 GHz, respectively.

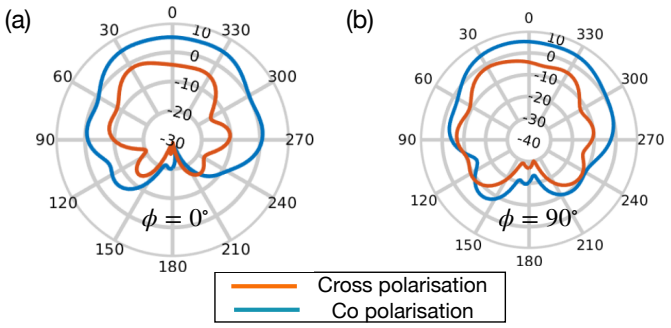

Fig. 13. Measurement of Co polarization and Cross polarization at f_1 = 3.75 GHz (a) Elevation radiation pattern ($\phi = 0°$), (b) Elevation radiation pattern ($\phi = 90°$)

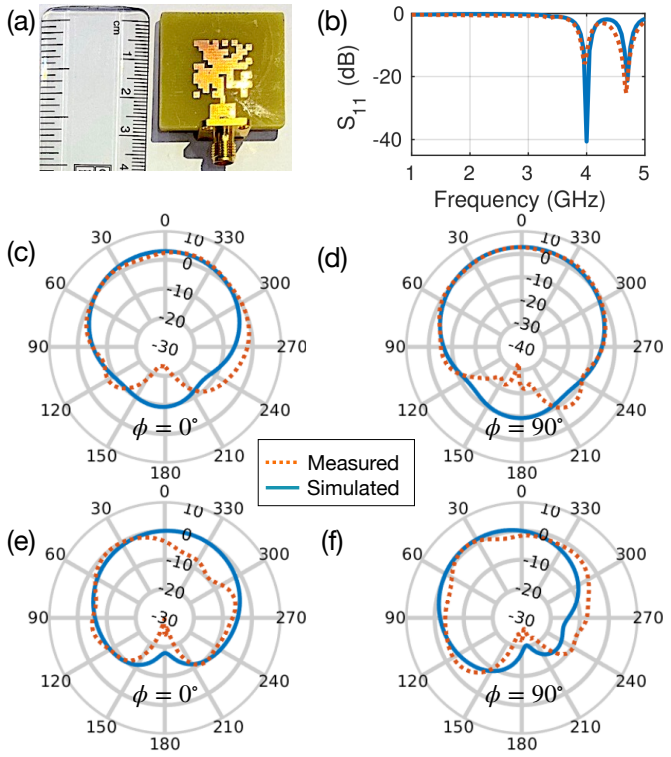

Fig. 14. (a) An instance of fabricated inverse designed dual-band antenna at f_1 = 4 GHz and f_2 = 4.7 GHz, (b) reconstructed spectrum vs the measured spectrum, Simulated and measured elevation radiation pattern at f_1 = 4 GHz (a) $\phi = 0°$, (b) $\phi = 90°$, at f_2 = 4.7 GHz (c) $\phi = 0°$, (d) $\phi = 90°$

VI. Conclusion

In this paper, we propose a data-efficient transfer learning based approach for the design of realistic antennas. Our approach builds on pre-trained neural networks for the design of other antennas, thereby making use of already acquired knowledge in order to reduce the dataset size requirements for learning a new task. Specifically, we solve the problem of designing multi-band antennas with a minimal number of dielectric simulations using knowledge transfer and a reduced network training time. To demonstrate the viability of the approach, we fabricate single and double band antennas which show excellent agreement between simulation and experimental results across RF (1-5 GHz) frequencies. We also perform a numerical sensitivity analysis to ascertain the robustness of

our design to fabrication issues. Our approach is fairly general and can be used to quickly design antennas on different substrates and at different frequency bands while economizing the computational burden of dataset generation. Complementary to conventional analytical or rule-based heuristic methods, such data-driven approaches can enable new functionalities in the field of high performance and multi-functional antenna design.

APPENDIX A
IMPROVED BPSO ALGORITHM

In this paper, we use the algorithm presented in [26] for inverse design. As proposed by the authors, this Improved Binary Particle Swarm Optimization (BPSO) algorithm addresses the limitations of the original BPSO method, and we opt for it due to its demonstrated superior convergence performance ([26], Fig. 8), compared to both the original BPSO [48] and another variant proposed in [49].

In our application, the algorithm takes a vector containing frequencies in the pass and stop bands, along with the proposed transfer learning based surrogate model as inputs and generates the optimal binary design as the output. The following equations governs the velocity, V_{ij}, where the subscript i, j refers to particle number i and the particular bit j of that particle's velocity, respectively, of each particle of the swarm in the Improved BPSO algorithm:

$$V_{ij}(t+1) = \begin{cases} \omega_t V_{ij}(t) + c_1 r_1 + c_2 r_2 & \text{if gBest = pBest = 1} \\ \omega_t V_{ij}(t) - c_1 r_1 - c_2 r_2 & \text{if gBest = pBest = 0} \\ \omega_t V_{ij}(t) & \text{otherwise,} \end{cases} \tag{5}$$

where ω_t is calculated as $(\frac{\omega_{max}-\omega_{min}}{N})t$ with ω_{max} and ω_{min} being the maximum and minimum values of ω, N is the maximum number of iterations and t is the current iteration. We chose ω_{max} = 0.9, ω_{min} = 0.4, c_1 = 2, c_2 = 2, N = 50, 100 (for single band and double band spectrums) and r_1 and r_2 are a uniformly distributed random number in the range (0,1).

The equations ensure that the particle's speed adjusts based on whether both (pBest and gBest) the best solutions agree. If they do, the speed changes. But if they don't agree, the speed remains the same, relying only on the previous speed. This is logical because if neither the overall best nor the personal best solution is favoured, there's no need for the speed to change. By using these equations, the authors in [26] discovered that the proposed algorithm converges more effectively.

ACKNOWLEDGEMENTS

The authors acknowledge M Sreenivasan and Prasad S for their valuable discussions and help conducting the radiation pattern experiments at SAMEER Perungudi, Chennai, India. Additionally, we thank Prof. Srujana Kagita, IIT Tirupati, for her guidance and insights into the work of these inverse-designed structures.

REFERENCES

[1] L.-Y. Xiao, W. Shao, F.-L. Jin, and B.-Z. Wang, "Multiparameter modeling with ann for antenna design," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 7, pp. 3718–3723, 2018.

[2] 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 Microwaves, Antennas & Propagation*, vol. 13, no. 2, pp. 149–155, 2019.

[3] Y. Sharma, H. H. Zhang, and H. Xin, "Machine learning techniques for optimizing design of double t-shaped monopole antenna," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 7, pp. 5658–5663, 2020.

[4] E. Zhu, Z. Wei, X. Xu, and W.-Y. Yin, "Fourier subspace-based deep learning method for inverse design of frequency selective surface," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 7, pp. 5130–5143, 2021.

[5] E. Zhu, E. Li, Z. Wei, and W.-Y. Yin, "Adversarial-network regularized inverse design of frequency-selective surface with frequency-temporal deep learning," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 10, pp. 9460–9469, 2022.

[6] M. Lamsalli, A. El Hamichi, M. Boussouis, N. Amar Touhami, and T. Elhamadi, "Genetic algorithm optimization for microstrip patch antenna miniaturization," *Progress In Electromagnetics Research*, vol. 60, pp. 113–120, 2016.

[7] A. Gupta, E. Karahan, C. Bhat, K. Sengupta, and U. K. Khankhoje, "Tandem neural network based design of multi-band antennas," *IEEE Transactions on Antennas and Propagation*, 2023.

[8] A. Gupta, C. Bhat, E. Karahan, K. Sengupta, and U. K. Khankhoje, "Machine learning based tandem network approach for antenna design," in *2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI)*, pp. 1–2, 2022.

[9] A. Gupta and U. K. Khankhoje, "Transfer learning based microstrip antenna design," in *2024 IEEE International Symposium on Antennas and Propagation and INC/USNC-URSI Radio Science Meeting (AP-S/INC-USNC-URSI)*, pp. 803–804, 2024.

[10] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics*, vol. 5, no. 4, pp. 1365–1369, 2018.

[11] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated nanophotonic devices," *arXiv preprint arXiv:1809.03555*, 2018.

[12] A. Gupta, U. K. Khankhoje, and C. Sideris, "Data-efficient and ultra-fast surrogate models for simulating nanophotonic power-splitters," in *2024 IEEE International Symposium on Antennas and Propagation and INC/USNC-URSI Radio Science Meeting (AP-S/INC-USNC-URSI)*, pp. 805–806, 2024.

[13] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, no. 1, pp. 1–40, 2016.

[14] R. Zhu, T. Qiu, J. Wang, S. Sui, C. Hao, T. Liu, Y. Li, M. Feng, A. Zhang, C.-W. Qiu, *et al.*, "Phase-to-pattern inverse design paradigm for fast realization of functional metasurfaces via transfer learning," *Nature communications*, vol. 12, no. 1, p. 2974, 2021.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[16] M. Kiani, M. Zolfaghari, and J. Kiani, "Transfer learning for inverse design of tunable graphene-based meta-surfaces," *Journal of Materials Science*, vol. 59, no. 8, pp. 3516–3530, 2024.

[17] E. A. Karahan, Z. Liu, and K. Sengupta, "Deep-learning-based inverse-designed millimeter-wave passives and power amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 58, no. 11, pp. 3074–3088, 2023.

[18] Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljacic, "Migrating knowledge between physical scenarios based on artificial neural networks," *ACS Photonics*, vol. 6, no. 5, pp. 1168–1174, 2019.

[19] Z. Fan, C. Qian, Y. Jia, M. Chen, J. Zhang, X. Cui, E.-P. Li, B. Zheng, T. Cai, and H. Chen, "Transfer-learning-assisted inverse metasurface design for 30% data savings," *Physical Review Applied*, vol. 18, no. 2, p. 024022, 2022.

[20] Y. Zhu, Y. Xie, H. Fan, C. Shen, Z. Chen, N. Bai, and X. Sun, "Inverse design of folded waveguide swss for application in twts based on transfer learning of deep neural network," *IEEE Transactions on Plasma Science*, vol. 50, no. 9, pp. 3276–3282, 2022.

[21] Y. Zhang, J. Xu, and S. Jiang, "Inverse design of tunable lowpass microstrip filters based on generative adversarial network and transfer learning," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 71, no. 5, pp. 2594–2598, 2024.

[22] F. Gao, Z. Ou, C. Yang, J. Yang, J. Deng, and B. Yan, "Cross-domain heterogeneous metasurface inverse design based on a transfer learning method," *Optics Letters*, vol. 49, no. 10, pp. 2693–2696, 2024.

[23] J. Zhang, Y.-D. Wang, Y. Wu, K. Kang, and W.-Y. Yin, "Inverse design of on-chip interconnect via transfer learning-based deep neural networks," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 13, no. 6, pp. 878–887, 2023.
[24] S. Bozinovski and A. Fulgosi, "The influence of pattern similarity and transfer learning upon training of a base perceptron b2," in *Proceedings of Symposium Informatica*, vol. 3, pp. 121–126, 1976.
[25] S. Bozinovski, "Reminder of the first paper on transfer learning in neural networks, 1976," *Informatica*, vol. 44, no. 3, 2020.
[26] 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," *Applied soft computing*, vol. 13, no. 12, pp. 4832–4840, 2013.
[27] C. A. Balanis, *Antenna theory: analysis and design*. John wiley & sons, 2015.
[28] T. Dozat, "Incorporating Nesterov Momentum into Adam," in *Proceedings of the 4th International Conference on Learning Representations*, pp. 1–4, 2016.
[29] L. Biewald, "Experiment tracking with weights and biases," 2020. Software available from wandb.com.
[30] Y. Bengio, I. Goodfellow, and A. Courville, *Deep learning*, vol. 1. MIT press Cambridge, MA, USA, 2017.
[31] N. Calik, F. Güneş, S. Koziel, A. Pietrenko-Dabrowska, M. A. Belen, and P. Mahouti, "Deep-learning-based precise characterization of microwave transistors using fully-automated regression surrogates," *Scientific reports*, vol. 13, no. 1, p. 1445, 2023.
[32] J. Wu, X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, and S.-H. Deng, "Hyperparameter optimization for machine learning models based on bayesian optimization," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, 2019.
[33] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, 2020.
[34] J. Mockus, "Application of bayesian approach to numerical methods of global and stochastic optimization," *Journal of Global Optimization*, vol. 4, pp. 347–365, 1994.
[35] M. F. A. Sree, M. H. Abd Elazeem, and W. Swelam, "Dual band patch antenna based on letter slotted dgs for 5g sub-6ghz application," in *Journal of Physics: Conference Series*, vol. 2128, p. 012008, IOP Publishing, 2021.
[36] Z. Akhter, R. M. Bilal, and A. Shamim, "A dual mode, thin and wideband mimo antenna system for seamless integration on uav," *IEEE Open Journal of Antennas and Propagation*, vol. 2, pp. 991–1000, 2021.
[37] A. G. Ambekar and A. A. Deshmukh, "E-shape microstrip antenna for dual frequency wlan application," *Progress In Electromagnetics Research C*, vol. 104, pp. 13–24, 2020.
[38] 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 *2013 Loughborough Antennas & Propagation Conference (LAPC)*, pp. 490–493, IEEE, 2013.
[39] H.-C. Tung and K.-L. Wong, "A shorted microstrip antenna for 2.4/5.2 ghz dual-band operation," *Microwave and Optical Technology Letters*, vol. 30, no. 6, pp. 401–402, 2001.
[40] S. K. Noor, M. Jusoh, T. Sabapathy, A. H. Rambe, H. Vettikalladi, A. M. Albishi, and M. Himdi, "A patch antenna with enhanced gain and bandwidth for sub-6 ghz and sub-7 ghz 5g wireless applications," *Electronics*, vol. 12, no. 12, p. 2555, 2023.
[41] S. Koziel, M. A. Belen, A. Çalişkan, and P. Mahouti, "Rapid design of 3d reflectarray antennas by inverse surrogate modeling and regularization," *IEEE Access*, vol. 11, pp. 24175–24184, 2023.
[42] C. A. Balanis, *Advanced engineering electromagnetics*. John Wiley & Sons, 2012.
[43] J. Eichler, P. Hazdra, M. Capek, and M. Mazanek, "Modal resonant frequencies and radiation quality factors of microstrip antennas," *International Journal of Antennas and Propagation*, vol. 2012, 2012.
[44] R. Garg, *Microstrip antenna design handbook*. Artech house, 2001.
[45] R. Lee, T. Huynh, and K. Lee, "Experimental study of the cross polarization characteristics of rectangular patch antennas," in *Digest on Antennas and Propagation Society International Symposium*, pp. 636–639, IEEE, 1989.
[46] M. Oberhart, Y. Lo, and R. Lee, "New simple feed network for an array module of four microstrip elements," *Electronics Letters*, vol. 9, no. 23, pp. 436–437, 1987.
[47] T. Huynh and K. Lee, "Cross polarization characteristics of rectangular patch antennas," in *1988 IEEE AP-S. International Symposium, Antennas and Propagation*, pp. 708–711, IEEE, 1988.
[48] J. Kennedy and R. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, vol. 5, pp. 4104–4108 vol.5, 1997.
[49] M. A. Khanesar, M. Teshnehlab, and M. A. Shoorehdeli, "A novel binary particle swarm optimization," in *2007 Mediterranean conference on control & automation*, pp. 1–6, IEEE, 2007.

Aggraj Gupta (Graduate Student Member, IEEE) received his B.E. degree from Panjab University, Chandigarh, and Masters from the Indian Institute of Technology (ISM) Dhanbad, in 2016 and 2018, respectively. He is currently a graduate student in the Department of Electrical Engineering at the Indian Institute of Technology Madras, Chennai.

Uday K Khankhoje (Senior Member, IEEE) received a B.Tech. degree in Electrical Engineering from the Indian Institute of Technology Bombay, Mumbai, in 2005, and the M.S. and Ph.D. degrees in Electrical Engineering from the California Institute of Technology, Pasadena, in 2010. He was a Caltech Postdoctoral Scholar at the Jet Propulsion Laboratory (NASA/Caltech) from 2011-12, a Postdoctoral Research Associate in the Department of Electrical Engineering at the University of Southern California, Los Angeles, USA, from 2012-13, and an Assistant Professor of Electrical Engineering at the Indian Institute of Technology Delhi, India from 2013-16. Since 2016 he has been at the Department of Electrical Engineering in the Indian Institute of Technology Madras, Chennai, where he is currently an Associate Professor and leads the numerical electromagnetics and optics (NEMO) research group which focuses on solving electromagnetics inspired inverse problems.