

Design-ready farfield pattern modeling of broadband antennas using multi-branch attention-based recurrent convolutional decoder neural network

Slawomir Koziel^{a,b,*}, Kaustab C. Sahu^a, Anna Pietrenko-Dabrowska^b

^a Engineering Optimization & Modeling Center, Reykjavik University, Reykjavik, 102, Iceland

^b Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Gdansk, 80-233, Poland

HIGHLIGHTS

- Comprehensive data-driven modeling of broadband antennas.
- Development of multi-branch recurrent convolutional decoder network.
- Utilization of attention mechanisms for reliable radiation pattern modeling.
- Competitive accuracy of the framework showcased against state-of-the-art benchmark.
- Real-world design applications demonstrated involving global/local surrogate optimization.

ARTICLE INFO

Keywords:

Broadband antennas
Neural networks
Deep learning
Data-driven modeling
Far-field pattern modeling
Design optimization

ABSTRACT

Behavioral modeling is becoming increasingly important in antenna development, particularly in accelerating simulation-based design processes such as optimization. However, the construction of accurate data-driven models remains a challenging task, typically limited to electrical responses such as reflection coefficients or integral performance metrics (e.g., gain). In contrast, the behavioral modeling of radiation patterns—although crucial from a design perspective—has been relatively scarce in the existing literature. This research introduces a new surrogate modeling framework for broadband antennas that leverages customized deep learning techniques. The proposed approach enables accurate and comprehensive modeling of both electrical responses (e.g., reflection coefficient, gain) and far-field radiation patterns across a broad range of frequencies, elevation/azimuth angles, and geometric parameters. As a result, the developed surrogates are highly versatile and applicable for addressing a wide spectrum of design and optimization tasks. For accurate representation of radiation patterns, a hybrid deep learning architecture is developed that integrates multi-branch recurrent processing with cross-attention-based feature fusion. Six independent RNN branches (each combining FC, LSTM, and GRU layers) extract complementary representations of the geometric input space, which are subsequently fused through a cascade of attention modules. The fused embedding is decoded using a transposed-convolutional refinement network to reconstruct the full radiation pattern. Extensive validation studies, performed using two ultra-wideband antenna designs, confirm the excellent predictive accuracy and robustness of the proposed surrogates. The results also demonstrate their advantages over multiple benchmark methods, including conventional regression models and various deep learning schemes. Furthermore, the practical utility of the framework is demonstrated through real-world design tasks, including impedance matching enhancement, size reduction, omnidirectional radiation, and gain stability. Overall, the presented methodology represents a significant advancement toward comprehensive multi-response antenna modeling and substantially extends the capabilities of behavioral metamodeling in contemporary antenna design practice.

* Corresponding author at: Engineering Optimization & Modeling Center, Reykjavik University, Reykjavik, 102, Iceland.

E-mail addresses: koziel@ru.is (S. Koziel), kaustab23@ru.is (K.C. Sahu), anna.dabrowska@pg.edu.pl (A. Pietrenko-Dabrowska).

<https://doi.org/10.1016/j.asoc.2026.116325>

Received 8 January 2026; Received in revised form 20 August 2026; Accepted 21 August 2026

Available online 11 September 2026

1568-4946/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The development of contemporary antenna systems is a complex, multi-objective endeavor. It must satisfy a wide range of performance requirements—such as high gain, broad bandwidth, and strong isolation [1–5], while also supporting advanced functionalities, including multi-band operation and circular polarization [6–9]. In addition, many designs are subject to stringent constraints related to miniaturization and compactness [10–14]. Meeting these diverse demands often leads to geometrically intricate antenna structures whose accurate evaluation and optimization depend heavily on electromagnetic (EM) analysis. EM simulation is indispensable for capturing essential electromagnetic effects such as mutual coupling, radiation losses, and the impact of extra components (e.g., mounting fixtures, radomes) [15–19]. Consequently, EM modeling is now integral to nearly all stages of antenna development, from architecture exploration and parametric studies to final optimization [20–23]. Among these stages, optimization remains the most computationally demanding [24][25], particularly when involving global search procedures [26–32] or multi-objective formulations [33–35]. The challenge becomes even more pronounced for population-based bio-inspired algorithms [36–45] including recently proposed evolutionary and fitness-dependent optimization schemes [46][47], which typically suffer from low computational efficiency [48–53]. Similarly, uncertainty quantification and yield-driven optimization introduce additional computational burdens, further emphasizing the need for efficient yet accurate design methodologies [54][55].

Accelerating EM-based antenna design undergoes extensive studies, leading to a wide range of methodologies aimed at reducing computational cost while maintaining accuracy. Among these are advanced gradient-based optimization schemes employing adjoint sensitivities [56][57], selective or restricted sensitivity updating strategies [58–60], mesh deformation techniques [61], and the use of fast EM solvers [62–64] or parallel computing frameworks [65]. Other conceptually distinct approaches include response feature-based optimization [66–68], cognition-driven design [69], and dimensionality reduction methods [70]. A particularly influential class of techniques is surrogate modeling, which has gained substantial traction in recent years for its ability to expedite EM-driven tasks such as global optimization [71][72], multi-objective optimization [73][74], and statistical or yield-driven design [75–77]. By replacing costly EM simulations with computationally inexpensive predictive models, surrogate-based methods can dramatically reduce overall design time. Among surrogate approaches, data-driven (behavioral) models, such as kriging, support vector regression, and artificial neural networks (ANNs) [78–82], are the most broadly adopted owing to their flexibility, straightforward implementation, and accessibility [83–86]. These metamodels are often integrated within machine learning (ML) frameworks, acting as fast predictors to generate and refine candidate designs [87–92]. In contrast, physics-based surrogates, though potentially more accurate, are typically problem-specific and primarily applied in local optimization contexts, including response correction and space mapping [93–96].

Recently, ANNs have attracted significant attention, particularly within the domain of machine learning (ML)-based optimization. Commonly employed network architectures include multilayer perceptrons (MLPs) [97–101], deep neural networks (DNNs) [102][103], convolutional NNs (CNNs) [104], and long short-term memory (LSTM) architectures [105]. In addition to these standard topologies, various custom network designs have been developed to address specific challenges, such as multi-fidelity networks, evolutionary ANNs, and autoencoders [106–108], demonstrating the effectiveness of neural networks for antenna-response modeling. Most ANN-based approaches to antenna modeling are designed as forward models that predict antenna performance from geometry and material parameters. Inverse modeling, which aims to infer geometry from desired performance characteristics, remains relatively uncommon [109–113] (see also [114]

for a recent review). More general-purpose ANN-based modeling frameworks are even rarer due to their inherent complexity and high data requirements. Notable recent developments in this area involve DNNs, CNNs, and graph neural networks (GNNs) [115–117], which have shown promise for capturing intricate electromagnetic behaviors across diverse antenna geometries. Nevertheless, these approaches were still directed towards electrical responses, inverse designs, or selected response quantities rather than a comprehensive frequency–angle radiation-pattern reconstruction.

Surrogate modeling for antennas, though highly beneficial, remains challenging. The primary difficulties arise from the curse of dimensionality, strong nonlinearity of system's responses, and the need for covering broad parameter and frequency ranges to make sure that the surrogate is suitable for practical design applications. Addressing these challenges is nontrivial. A commonly adopted strategy, particularly in optimization-oriented workflows, is utilization of machine learning (ML) frameworks, where metamodels are constructed only within limited, promising subregions of the design space determined early during the search run [118]. However, this approach leaves large portions of the space unexplored. Several other methodologies have been developed to alleviate the curse of dimensionality, including principal component analysis (PCA), least angle regression (LAR), and high-dimensional model representation (HDMR) [119–121], as well as multi-fidelity modeling techniques [122][123]. More recently, performance-driven modeling [124] has emerged as an effective solution, improving modeling efficiency by focusing surrogate construction on carefully defined regions that contain high-quality designs. Nevertheless, determining these regions entails substantial initial computational overhead, as it typically requires pre-optimizing reference designs [125] or pre-screening the parameter space [126]. This added cost can be mitigated through variable-fidelity electromagnetic (EM) analysis [127], incorporation of sensitivity information (e.g., gradient-enhanced kriging) [128], or the use of advanced deep learning architectures that can capture complex parameter interactions more efficiently [129]; highlighting the limitations of neural networks depending on their ability to represent complex EM-response structures under limited training-data conditions.

Most existing data-driven antenna modeling techniques are confined to representing vector-valued responses, such as electrical characteristics (e.g., reflection coefficients versus frequency) or integral parameters (e.g., gain versus frequency). In contrast, modeling of far-field radiation patterns—crucial for tasks like beam shaping, sidelobe suppression, front-to-back ratio improvement, or achieving omnidirectional radiation—has received limited attention. The distinction is important, since radiation patterns are not merely simple vector-valued outputs; they are structured responses that vary jointly with frequency and angular direction. A notable exception is array-factor (AF)-based modeling, commonly used for antenna arrays. However, AF models suffer from inherent limitations, primarily due to the neglect of mutual coupling effects. These effects can be incorporated through simulation-based superposition models [17] or EM analysis of the entire array. To enhance AF modeling, several studies have integrated AF representations with different ANN architectures, including MLPs, graph neural networks (GNNs), and CNNs, for behavioral modeling of array radiation patterns [117][130–134]. Recent works have attempted to incorporate radiation pattern modeling directly into surrogate-based design frameworks. For example, CNNs have been employed to represent far-field responses as surrogate outputs in multi-objective optimization [135], while other studies have used radiation patterns as model inputs for inverse design, leveraging MLPs [116] or physics-informed neural networks [109]. Furthermore, adaptive neuro-fuzzy inference systems (ANFIS) have been applied to specific antenna types such as circular loop and corrugated horn antennas [136][137]. In another example, CNNs were used to model the near-field patterns of a simple patch antenna characterized by four geometric parameters [115]. Meanwhile, [138] proposed an autoencoder with convolutional layers to reconstruct

Symbol	Explanation		
$\mathbf{x} = [x_1 \dots x_n]^T$	Design variables (typically, antenna dimensions in mm)		
X	Parameter space delimited by lower and upper variable bounds $\mathbf{l} = [l_1 \dots l_n]^T$ and $\mathbf{u} = [u_1 \dots u_n]^T$		
$\mathbf{R}_{EM}(\mathbf{x})$	(Generic) responses of the computational (EM) simulation model of the antenna		
$S_{11}(\mathbf{x}, f)$	Specific responses: reflection as a function of design vector $\mathbf{x}$ and frequency f		
$G(\mathbf{x}, f)$	Specific responses: realized gain as a function of design vector $\mathbf{x}$ and frequency f		
$P(\mathbf{x}, f, \theta)$	Specific responses: radiation pattern as a function of design vector $\mathbf{x}$ and frequency f , and azimuth angle θ (here, two types of patterns are considered, H- and E-plane)		
$\mathbf{R}_s(\mathbf{x})$	(Generic) data-driven surrogate model response to be constructed		

(a)

Surrogate model type	Inputs		Output
	Type	Ranges	
$S_{11,s}(\mathbf{x}, f)$	Design variables $\mathbf{x}$	$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$ (parameter space X)	Complex reflection coefficient
	Frequency f	$f_L \leq f \leq f_H$ (frequency range F)	
$G_s(\mathbf{x}, f)$	Design variables $\mathbf{x}$	$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$ (parameter space X)	Realized gain in broadside direction
	Frequency f	$f_L \leq f \leq f_H$ (frequency range F)	
$P_{H,s}(\mathbf{x}, f, \theta)$ $P_{E,s}(\mathbf{x}, f, \theta)$	Design variables $\mathbf{x}$	$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$ (parameter space X)	H-plane and E-plane radiation patterns
	Frequency f	$f_L \leq f \leq f_H$ (frequency range F)	
	Azimuth angle θ	$\theta_L \leq \theta \leq \theta_H$ (angle range Θ)	

(b)

Fig. 1. Comprehensive antenna modeling: (a) basic terminology, (b) modeled characteristics.

3D radiation patterns from two-dimensional field cuts. Only a few studies have pursued comprehensive modeling of antenna characteristics that include far-field responses [139–141]. For instance, [140] utilized MLPs for inverse modeling in a multi-objective optimization framework, albeit over limited parameter ranges. Similarly, [141] demonstrated a hybrid approach combining MLPs with support vector regression (SVR) for modeling a Fabry–Pérot antenna, though the setup was relatively simple (involving just three design variables). Importantly, most of these works lack rigorous error analysis and benchmarking, which limits their generalizability and comparative assessment.

The objective of this study is to develop a unified data-driven surrogate framework capable of high-fidelity modeling of both electrical characteristics (such as reflection and gain) and far-field responses across a broad frequency range and wide elevation/azimuth angles, while accounting for variations in antenna geometry. Within this framework, the primary focus and principal contribution of the paper lie in modeling radiation patterns, for which a customized neural architecture is introduced. Specifically, the radiation-pattern prediction task is addressed through a tailored attention–fusion network (AFN) that employs a multi-stream recurrent architecture with cross-attention, coupled with a hierarchical decoder. The network is designed to process compact geometric design parameters through multiple parallel recurrent branches composed of fully connected, long short-term memory (LSTM), and gated recurrent unit (GRU) layers. The latent representations produced by the parallel branches are then fed into a set of cross-attention modules, which learn to selectively emphasize informative branches while suppressing redundant or noisy ones. The fused output is subsequently reshaped and processed by a hierarchical transposed-convolutional decoder that reconstructs the full radiation pattern. This design enables the attention-fusion net (AFN) to outperform conventional CNN, RNN, and transformer baselines by explicitly

modeling inter-branch interactions while preserving fine-scale angular–frequency structure. Comprehensive verification studies were performed using two broadband antennas with the frequency range from 3.1 GHz to 10.6 GHz. The results demonstrate the exceptional accuracy and robust generalization capability of the presented surrogates, outperforming a wide range of benchmark techniques, including traditional regression models and modern deep learning architectures such as MLPs, CNNs, and RNNs. The error analysis was further complemented with practical design case studies, including antenna optimization for impedance matching, size minimization, omnidirectional radiation, and gain stability. In all scenarios, the proposed models consistently delivered reliable predictions and facilitated the identification of high-performance designs, which were subsequently validated through EM simulations. The presented methodology may be considered a step towards comprehensive multi-response antenna modeling, which extends the capabilities of behavioral metamodels in contemporary antenna development.

This work delivers several key original contributions:

- An innovative deep learning-oriented framework for holistic modeling of antenna electrical and field characteristics across wide ranges of frequency elevation/azimuth angles, and diverse geometric configurations.
- Development of a multi-branch parallel recurrent pipeline (with fully connected - LSTM - GRU layers) designed to learn complementary nonlinear transformations of the geometric parameter space.
- Development of a cross-attention-based fusion mechanism that dynamically weights representations to enable the network to model inter-branch dependencies.
- Development of a hierarchical transposed-convolutional decoder that reconstructs full high-resolution radiation patterns from the fused latent embedding, preserving its structure.

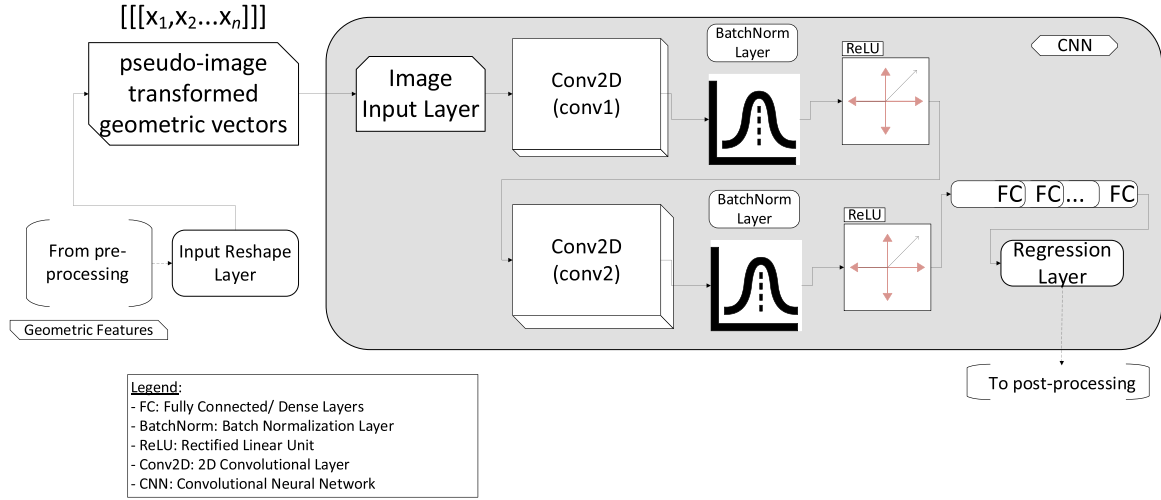


Fig. 2. CNN used to represent antenna's reflection and gain characteristics.

- Comprehensive performance evaluation underscoring the proposed procedure's exquisite predictive power and cost efficiency over a broad spectrum of benchmark modeling approaches.
- Application of the models to practical design optimization problems, highlighting their effectiveness in scenarios such as impedance matching enhancement, size reduction, and omnidirectional radiation characteristics.

Overall, the main difference between the proposed model and previous works is that the present approach is specifically designed for structured radiation-pattern surrogate modeling, rather than only vector-valued electrical-response prediction or isolated field-reconstruction tasks. The proposed AFN combines recurrent feature extraction, attention-based fusion, and convolutional decoding to reconstruct frequency–angle radiation-pattern responses. This distinguishes the method from prior approaches that mainly focus on conventional vector-valued response prediction or simpler field-modeling settings.

2. Modelling of antenna responses using multi-branch recurrent convolutional decoder network

This part of the paper elaborates on the proposed antenna modeling methodology. The modeling task is formally stated in Section 2.1. The architecture and other relevant details of neural networks developed to represent antenna reflection, gain, and far-field radiation patterns are discussed in Sections 2.2 and 2.3, respectively. The complete modeling procedure is put together in Section 2.4.

2.1. Surrogate modeling of antenna characteristics: problem formulation

In this research, we construct surrogate models representing the electrical and field responses of broadband antennas. More specifically, we are interested in the reflection coefficient and realized gain versus frequency, as well as the H-plane and E-plane radiation patterns versus frequency f and azimuth angle θ . The notation used in this context is explained in Fig. 1. The regions of validity of the specific surrogates are $X \times F$ (for $S_{11,s}(x,f)$ and $G_s(x,f)$), and $X \times F \times \Theta$ for $P_s(x,f,\theta)$. The angle range Θ for radiation pattern modeling is from $\theta_L = -180^\circ$ to $\theta_H = +180^\circ$. The frequency and angle ranges are discretized: we use M_f frequency samples f_k , $k = 1, \dots, M_f$, uniformly allocated in F , and M_θ angles θ_k , $k = 1, \dots, M_\theta$, uniformly allocated within Θ . Typically, M_f is about 20, whereas $M_\theta = 181$ (corresponding to an angle step of one degree).

The model's reliability is evaluated using a relative root-mean-square error (RRMSE), defined as $\|R_s(x) - R_f(x)\| / \|R_f(x)\|$, where $\|$

$\|$ is a Frobenius norm. In the verification part of the paper, RRMSE will be computed using an independent set of 200 testing points uniformly distributed in the modeling domain. Note that the relative error is well-compatible with the visual assessment of EM-simulated and model-predicted outputs. In particular, good visual alignment (which is imperative for making the surrogate suitable for design purposes) typically corresponds to an RRMSE of a few percent. Alternative error metrics employed in engineering practice can be found, for example, in [142] and [143].

2.2. Convolutional Neural Network (CNN) models for electrical response modelling

To accurately model the electrical responses of the broadband antennas, specifically, reflection and gain characteristics, we employ a lightweight Convolutional Neural Network (CNN). This network processes normalized, image-like representations of the geometric parameters and predicts the corresponding electric response curves.

2.2.1. Model setup/pre-processing

The geometric parameters are first normalized and reshaped into a pseudo-image of dimensions $[1 \times 1 \times F]$, where F is the number of geometric features. This representation allows us to use efficient convolutional operations while preserving the original feature ordering.

2.2.2. Architecture

The model begins with an ImageInputLayer receiving the pseudo-image representation. This is followed by a sequence of 2D convolutional layers with increasing channel widths ($N \rightarrow 2N \rightarrow 4N$) and kernel sizes ($1 \times 1 \rightarrow 2 \times 1 \rightarrow 3 \times 1$). Convolutional layers are paired with batch normalization and ReLU activation to regularize training and provide non-linearity. After the convolutional blocks, a fully connected layer transforms the learned feature maps to a compact latent representation, which then feeds into a final regression layer to output the predicted responses. Fig. 2 shows the CNN model's flow diagram.

2.2.3. Rationale and Design Motivation

The model architecture is chosen for its good balance between complexity, computational efficiency, and generalization. The design incorporates several choices tailored to the structure of the current problem:

- Feature mixing using 1×1 convolutions – Conceptually inspired by Google's Inception module [144], the initial convolution performs

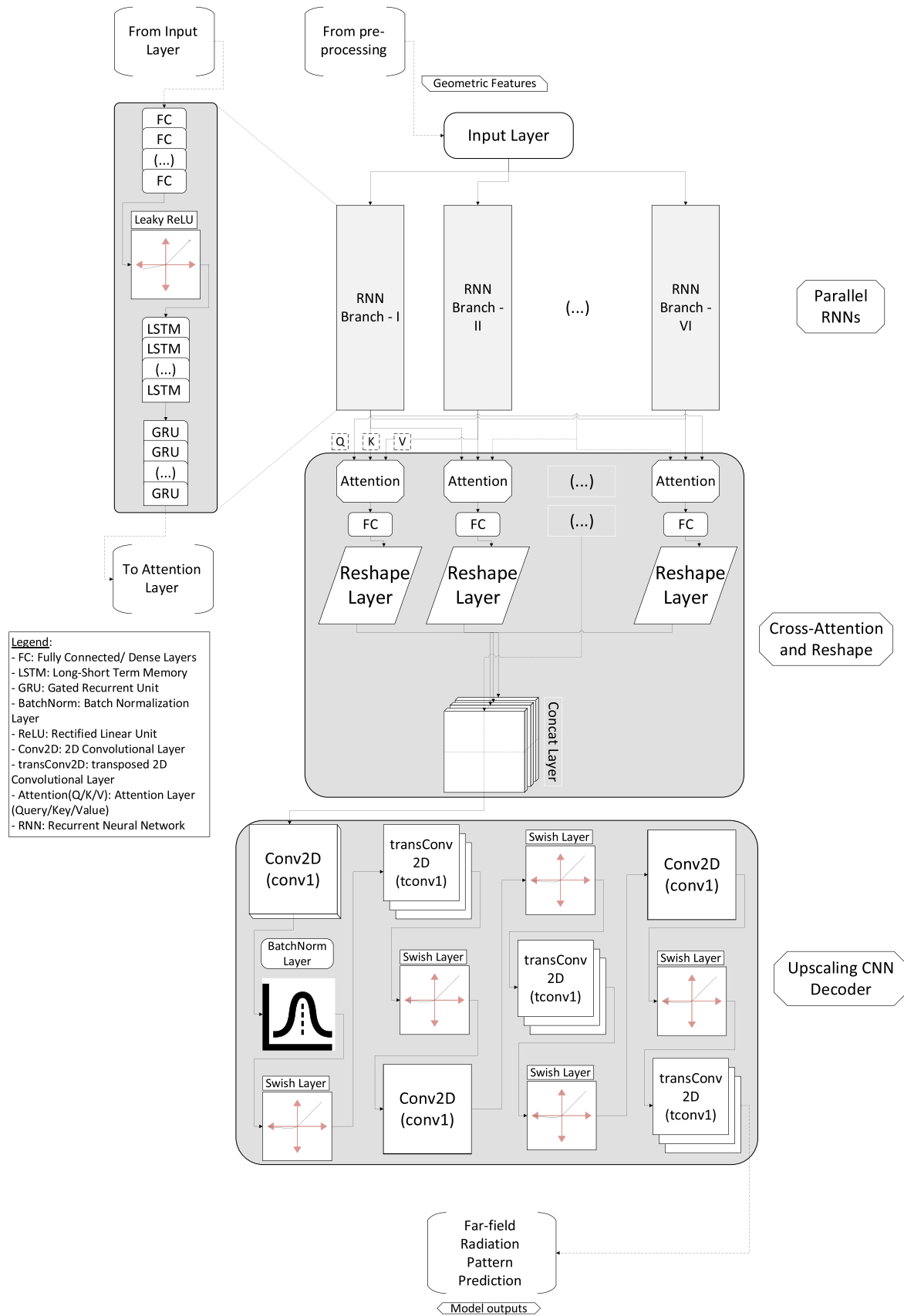


Fig. 3. Multi-branch attention-based recurrent convolutional decoder network for far-field radiation pattern modeling.

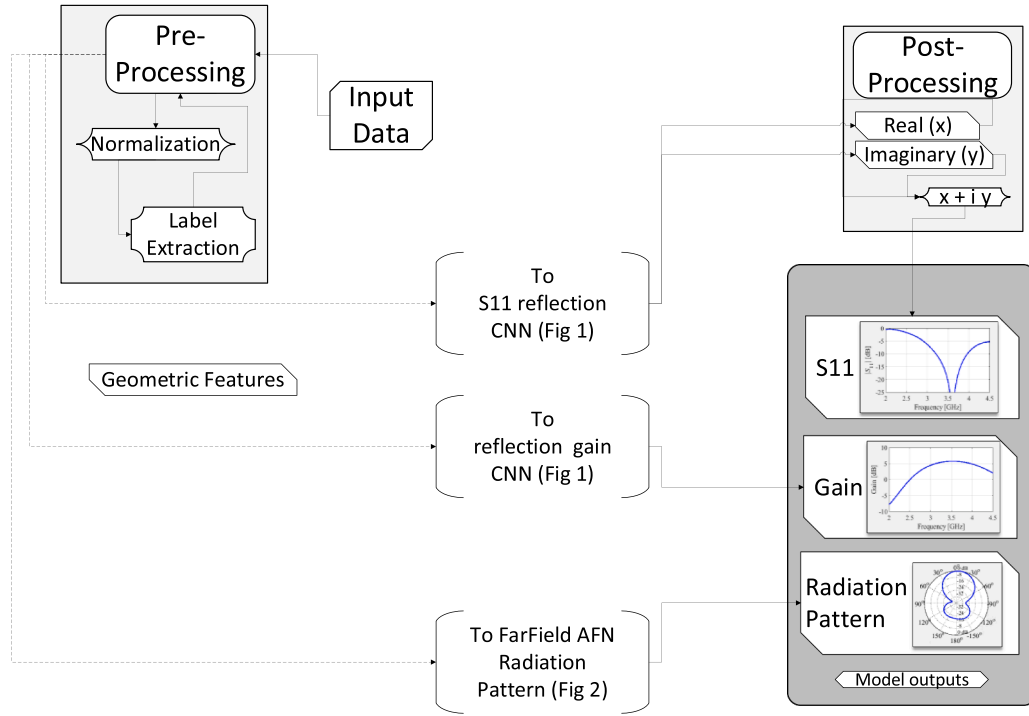


Fig. 4. Overall topology and data flow of the suggested CNN/AFN-based antenna model.

Table 1
Convolutional/Transposed-Convolutional Layers of the Decoder.

Block Type	Function	Kernel	Channels	Stride
Convolutional (conv2D)	Reduction	3×3	64	(2,1)
Transposed Conv. (transposedConv2D)	Upsampling - 1	3×3	64	(7,2)
Convolutional (conv2D)	Refinement	3×3	128	(1,1)
Transposed Conv. (transposedConv2D)	Upsampling - 2	3×3	64	(1,2)
Convolutional (conv2D)	Final Refinement	3×3	32	(1,1)
Transposed Conv. (transposedConv2D)	Output	3×3	1	(1,1)
	Reconstruction			

Table 2
Tested hyperparameter ranges.

Hyperparameter	Tested range/values
Batch size	8 - 32
Initial learning rate	7.5×10^{-5} to 2.5×10^{-3}
FC projection size	16 - 64
LSTM units	32 - 128
GRU units	64 - 128
Attention dimension	64 - 128
Decoder channel settings	Filter combinations ranging from 16 to 256 filters per layer.
Number of branches	4, 6
Number of epochs	128, 256

lightweight channel mixing. This acts as a learned linear projection across feature channels while keeping the costs low.

- Hierarchical receptive field expansion – the subsequent 2×1 and 3×1 kernels allow the model to progressively capture pairwise and multi-feature interactions, which often underlie EM interactions. This creates an inductive bias well-matched to the antenna behavior, where the features tend to act jointly rather than independently.
- Regularization through batch normalization – Each convolutional block includes batch normalization to prevent overfitting and to allow a smooth mapping from geometry to response. As shown in

Table 3
Hyperparameter setting for training of the proposed CNN/AFN models.

Hyperparameter	Electrical Response CNN	Radiation pattern AFN
Batch size	16	
Number of epochs	256	
Initial learning rate	$7.50E-04$	
Hidden dimensions	- LSTM layers: 64 units - GRU layers: 96 units	- Convolutional filters: progressively increased from 64 to 256
Optimizer	Adam [147]	

Section 3, overfitting was a recurring problem with several benchmark models.

In summary, this architecture is selected because it introduces the appropriate inductive bias for antenna modelling—namely, capturing localized cross-feature interactions—while maintaining low computational cost and strong generalization. Although RNNs and Transformers can emulate interaction patterns through recurrent or attention-based mechanisms, they tend to overfit when trained on limited data, requiring substantially larger parameter budgets and training resources.

2.3. Multi-branch attention-based recurrent convolutional decoder network for far-field radiation pattern modeling

To model the full 2D far-field radiation pattern from geometric design parameters, we propose an extensively customized multi-branch recurrent attention architecture followed by a convolutional decoder. This design is inspired by transformer-style attention mechanisms [145] and the idea of learning complementary representations through multi-branch processing, similar to Mixture-of-Experts (MoE) models [146], although our formulation differs substantially from both. Each branch extracts a distinct latent representation of the input parameters, and the cross-branch attention modules selectively fuse these

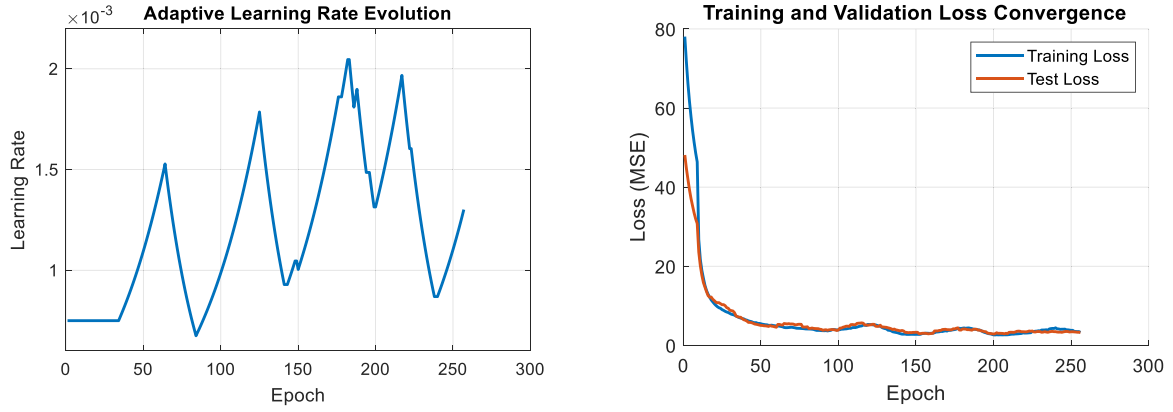


Fig. 5. Learn rate evolution and typical loss function convergence for the proposed CNN/AFN-based surrogate model.

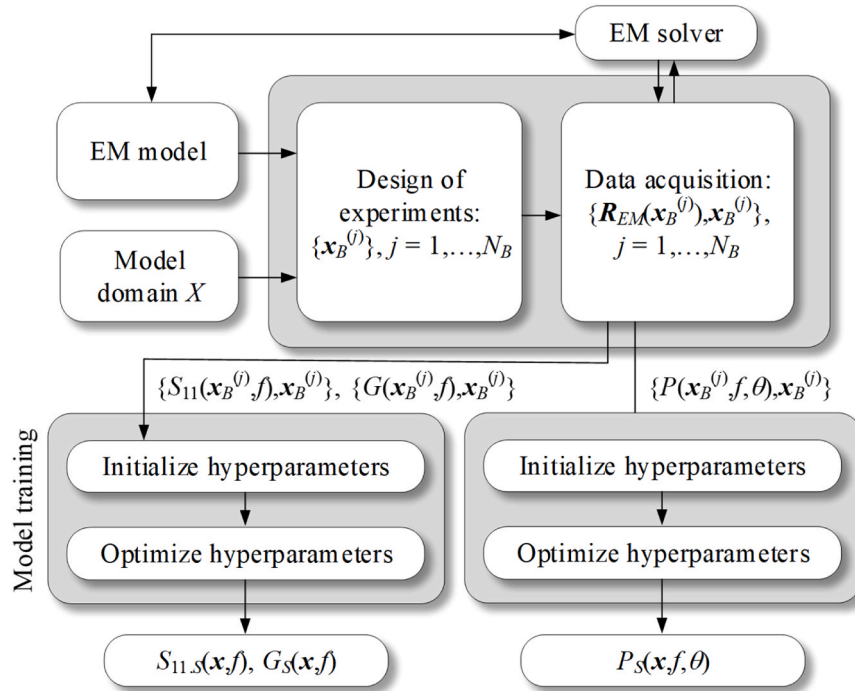


Fig. 6. Surrogate modeling of antennas using an attention-based recurrent convolutional decoder NN: complete procedure.

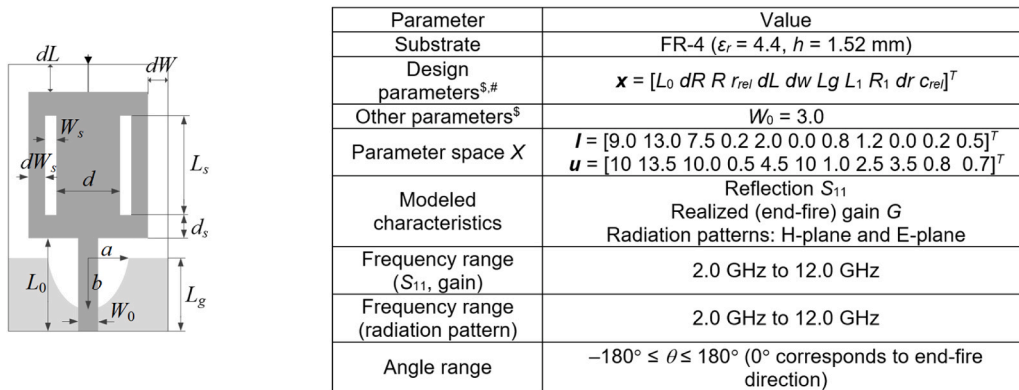
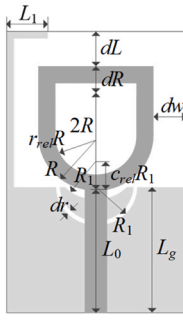


Fig. 7. Antenna II: miniaturized ultra-wideband monopole with elliptical ground-plane slot [149].

representations into a compact set of feature tokens. These fused tokens are then concatenated and passed into a CNN-based decoder, which

employs a sequence of transposed convolution (upsampling) layers with 3×3 kernels and unit stride (with appropriate padding) to



Parameter	Value
Substrate	RF-35 ($\epsilon_r = 3.5$, $h = 0.762$ mm)
Design parameters ^{s,#}	$\mathbf{x} = [L_0 \ dR \ R \ r_{rel} \ dL \ dw \ L_g \ L_1 \ R_1 \ dr \ c_{rel}]^T$
Other parameters ^s	$w_0 = 1.7$
Parameter space X	$\mathbf{l} = [8.0 \ 0.5 \ 4.0 \ 0.5 \ 3.0 \ 2.5 \ 9.0 \ 4.5 \ 2.0 \ 0.7 \ 0.4]^T$ $\mathbf{u} = [10.0 \ 1.2 \ 6.0 \ 0.7 \ 4.0 \ 4.5 \ 10.0 \ 5.5 \ 3.0 \ 1.0 \ 0.7]^T$
Modeled characteristics	Reflection S_{11} Realized (end-fire) gain G Radiation patterns: H-plane and E-plane
Frequency range (S_{11} , gain)	2.0 GHz to 12.0 GHz
Frequency range (radiation pattern)	2.0 GHz to 12.0 GHz
Angle range	$-180^\circ \leq \theta \leq 180^\circ$ (0° corresponds to end-fire direction)

Fig. 8. Antenna I: miniaturized ultra-wideband monopole with modified ground and L-shaped ground plane stub [150].

Table 4

Benchmark Methods.

Method	Setup
Linear Regression (LR) [151]	- Separate linear regression models per output dimension - Least squares learner
Gaussian process regression (GPR) [152]	- Rational Quadratic kernel functions - Separate GPR models for different responses - Hyperparameters optimized through maximum likelihood estimation
Support vector regression (SVR) [153]	- Gaussian(RBF) kernel function - Separate SVM models for different parts of the response - Kernel scale optimized automatically for each frequency
Feedforward Neural Network 1 (FF-NN) [154]	- Pyramidal Feedforward neural network with 4 hidden layers ($512 \rightarrow 256 \rightarrow 128 \rightarrow 64$ neurons) - ReLU activation functions between layers - Trained using Adam optimizer with a learning rate of 7.5×10^{-4}, 256 epochs
Recurrent Neural Network (RNN_1) [155]	- RNN with LSTM, GRU, and BiLSTM layers: LSTM (256 units), GRU (512 units), BiLSTM (512 units) - Fully connected layer at the end with regression output - Trained using Adam optimizer with a learning rate of 7.5×10^{-4}, 256 epochs
Recurrent Neural Network (RNN_2) [156]	- Attention-based Recurrent Neural Network with LSTM (256 units) and GRU (512 units) layers - Attention mechanism with two layers (512 units each) - Fully connected output layer with regression output, using Adam optimizer and a learning rate of 7.5×10^{-4}, 256 epochs
Convolutional Neural Network (CNN) [156]	- Convolutional network with two 2D convolutional layers (32 and 64 filters), followed by batch normalization and ReLU activation - Fully connected layers leading to regression output - Trained using Adam optimizer with a learning rate of 7.5×10^{-4}, 256 epochs
Transformers Encoder [156]	- Transformer model with stacked encoder modules: 4 blocks of Transformer encoders, each with 128 self-attention and feed-forward layers, using 8 attention heads, a dropout rate of 0.05 - Adaptive learning rate adjustment, using Adam optimizer and a learning rate of 7.5×10^{-4}, 256 epochs - Supports early stopping and checkpoint saving, with real-time training and validation loss plotting.

progressively lift the latent features into a smooth, high-resolution 2D radiation pattern. Its specifications are diagrammatically represented in Fig. 3.

2.3.1. Network specifications

The presented model architecture transforms geometric design parameters into a high-resolution 2D far-field radiation pattern through a series of structured processing blocks. The three main stages can be distinguished: (i) parallel recurrent branches to construct complementary latent representations, (ii) a cross-attention fusion module, and finally (iii) a convolutional decoder that upsamples the fused representations into the final spatial radiation pattern. The stages are visualized in Fig. 3 (see also Fig. 4 for the overall architecture), and a brief description is presented as follows:

Parallel Representations: Inspired by the Mixture-of-Experts (MoE) models [146], we design a parallel processing unit that receives the same input but processes it independently to generate distinct vector embeddings. A branch begins with a fully connected projection layer mapping the geometric parameters to a higher-dimensional latent space. The next block is a leaky ReLU activation, which helps stabilize gradients and avoid saturation effects. These projection features are then treated as a 1D sequence and passed to the recurrent stack of LSTM and GRU layers. Although the input geometric parameters are not inherently sequential, the fully-connected projection layer maps the design vectors into a higher-dimensional pseudo-sequential latent representation. This allows the recurrent units to model ordered dependencies among the geometric features through the projection latent. The LSTM focuses on capturing longer-range dependencies, while the subsequent GRU provides a lighter, complementary filtering stage that refines the latent state. This cascade balances representational power with parameter efficiency, reducing the possibility of overfitting that would result from using a deeper single-type recurrent block.

Cross-Branch Attention Fusion: The latent representations from the branches are fed into a set of cross-attention modules, enabling each branch to interrogate the information extracted by the others. For each attention block, the latent vector of one branch is used to form the query, while the remaining branch outputs serve as keys and values. This cross-branch configuration enables the model to selectively integrate features that recur across branches and suppress representations that are unique but uninformative. The output of this module – the attention refined embeddings are fused to form the multi-channel latent representation that initializes the decoding stage.

Convolutional Decoder for Spatial Reconstruction: The decoder maps the fused latent representation into a two-dimensional radiation grid through a sequence of transposed convolution (upsampling) blocks. The first layer reshapes the concatenated tokens into a low-resolution spatial tensor, establishing an initial geometric structure from which the full pattern is constructed. Subsequent decoder blocks consist of transposed convolutions, combined with appropriate padding, to ensure smooth and stable enlargement of the spatial resolution. Furthermore, batch normalization is applied to maintain stable feature statistics during upsampling, and ReLU activations inject the necessary non-linearity without distorting the regression output space. Skip connections were

Table 5
Radiation Pattern Modeling Results for Antenna I.

Modeling method	RRMSE [%]								
	$N_B = 500$			$N_B = 1000$			$N_B = 2000$		
	H-plane pattern	E-plane pattern	Mean [#]	H-plane pattern	E-plane pattern	Mean [#]	H-plane pattern	E-plane pattern	Mean [#]
LR	8.3	7.9	8.1	7.0	6.6	6.8	5.2	4.8	5.0
GPR	4.3	4.1	4.2	3.4	3.2	3.3	2.5	2.6	2.5
SVR	4.9	4.0	4.5	3.8	3.1	3.5	2.6	2.8	2.7
FF-NN	7.9	9.7	8.8	6.5	7.6	7.1	4.4	4.7	4.6
RNN_1	9.5	10.6	10.1	6.8	7.5	7.2	3.5	3.8	3.7
RNN_2	9.8	9.3	9.5	7.1	6.8	7.0	3.7	3.6	3.7
CNN	7.2	5.6	6.4	5.8	5.4	3.7	3.8	3.8	3.7
Transformers	3.8	4.7	4.2	3.2	3.6	2.3	2.7	2.5	2.5
This work	3.6	2.9	3.3	2.3	1.9	2.1	0.9	0.9	0.9

RRMSE averaged for H-plane and E-plane patterns.

Table 6
Radiation Pattern Modeling Results for Antenna II.

Modeling method	RRMSE [%]								
	$N_B = 500$			$N_B = 1000$			$N_B = 2000$		
	H-plane pattern	E-plane pattern	Mean [#]	H-plane pattern	E-plane pattern	Mean [#]	H-plane pattern	E-plane pattern	Mean [#]
LR	13.2	15.4	14.3	11.1	10.5	10.8	7.9	6.4	7.1
GPR	8.1	6.6	7.4	6.0	4.5	5.3	3.3	2.7	3.0
SVR	8.6	6.0	7.3	6.2	4.6	5.4	3.2	2.8	3.0
FF-NN	13.0	13.8	13.4	11.7	12	11.9	9.5	9.1	9.3
RNN_1	17.5	15.2	16.3	11.6	10.8	11.2	5.1	5.4	5.3
RNN_2	17.5	14.9	16.2	13.6	11.6	12.6	8.3	7.0	7.7
CNN	7.0	6.9	7.0	6.0	5.9	4.3	4.0	4.2	4.2
Transformers	9.0	6.8	7.9	8.1	7.0	6.2	4.3	5.3	5.3
This work	2.2	1.9	2.0	2.1	1.9	2.0	2.0	1.7	1.8

RRMSE averaged for H-plane and E-plane patterns.

not used, as the sequential decoder already provided accurate reconstruction while maintaining compact architecture. The decoder layers are schematically illustrated in Fig. 3 and described in detail in Table 1.

2.3.2. Design motivation and rationale

The architecture is designed to address the following modeling challenges: automatic feature relevance and accurately modeling radiation patterns in a low-data environment, and most importantly, recovering a high-resolution spatial field from a compact geometric descriptor. Therefore, we incorporated several design choices tailored to the current problem:

- Parallel recurrent branches mitigate the risk of over-specialization by allowing multiple latent hypotheses to form independently from the same input. Furthermore, the LSTM-GRU cascade enables each branch to capture both long- and short-range interactions among geometric parameters, which is particularly suitable for highly nonlinear interactions.
- A cross-attention model to integrate only the most mutually informative components across branches.
- Finally, the convolutional decoder is motivated by the spatial structure of radiation patterns.

Overall, the architecture integrates parallel recurrent encoding, attention-based consolidation, and convolutional spatial modeling to produce a deep, highly specialized network tailored for radiation pattern reconstruction.

2.4. Model training and hyperparameter configuration

This section outlines the training procedure and hyperparameter configurations used for the proposed model, covering both electrical response prediction and far-field radiation pattern representation. All

models were implemented in MATLAB using the Deep Learning Toolbox with custom training loops. To improve reproducibility, the tested ranges for the main hyperparameters are reported in Table 2, while the final selected hyperparameter values employed in each model are summarized in Table 3. An adaptive learning rate strategy is employed to promote stable convergence during training. The validation loss is monitored using a fixed sliding window of 32 epochs, and the mean loss over the current window is compared against that of the preceding window to assess convergence trends. If the relative improvement falls below 0.01%, the learning rate is reduced to 96% of its current value to facilitate finer-grained optimization. Conversely, sustained improvement beyond this threshold triggers a modest increase to 102.4% of the current learning rate, encouraging faster convergence. The resulting evolution of the learning rate over training is illustrated in Fig. 5.

The selected hyperparameter values were determined through a combination of preliminary exploratory experiments and empirical refinement across multiple training runs. In particular, the batch size, hidden dimensionality, and initial learning rate were chosen to balance convergence stability, representational capacity, and computational efficiency, while mitigating overfitting observed in deeper or wider configurations. Once a stable operating regime was identified, these hyperparameters were fixed and applied consistently across all reported experiments, including benchmark models, to ensure a fair and controlled comparison.

2.5. Complete procedure

Here, we summarize the entire modeling algorithm. As mentioned, the metamodel is constructed to predict the reflection coefficient and realized gain (both as a function of frequency), and H-plane and E-plane patterns as functions of frequency and azimuth angle from -180° to $+180^\circ$. The relevant terminology has been explained in Section 2.1. The surrogate is built based on training points, $\mathbf{x}_B^{(j)}, j = 1, \dots, N_B$,

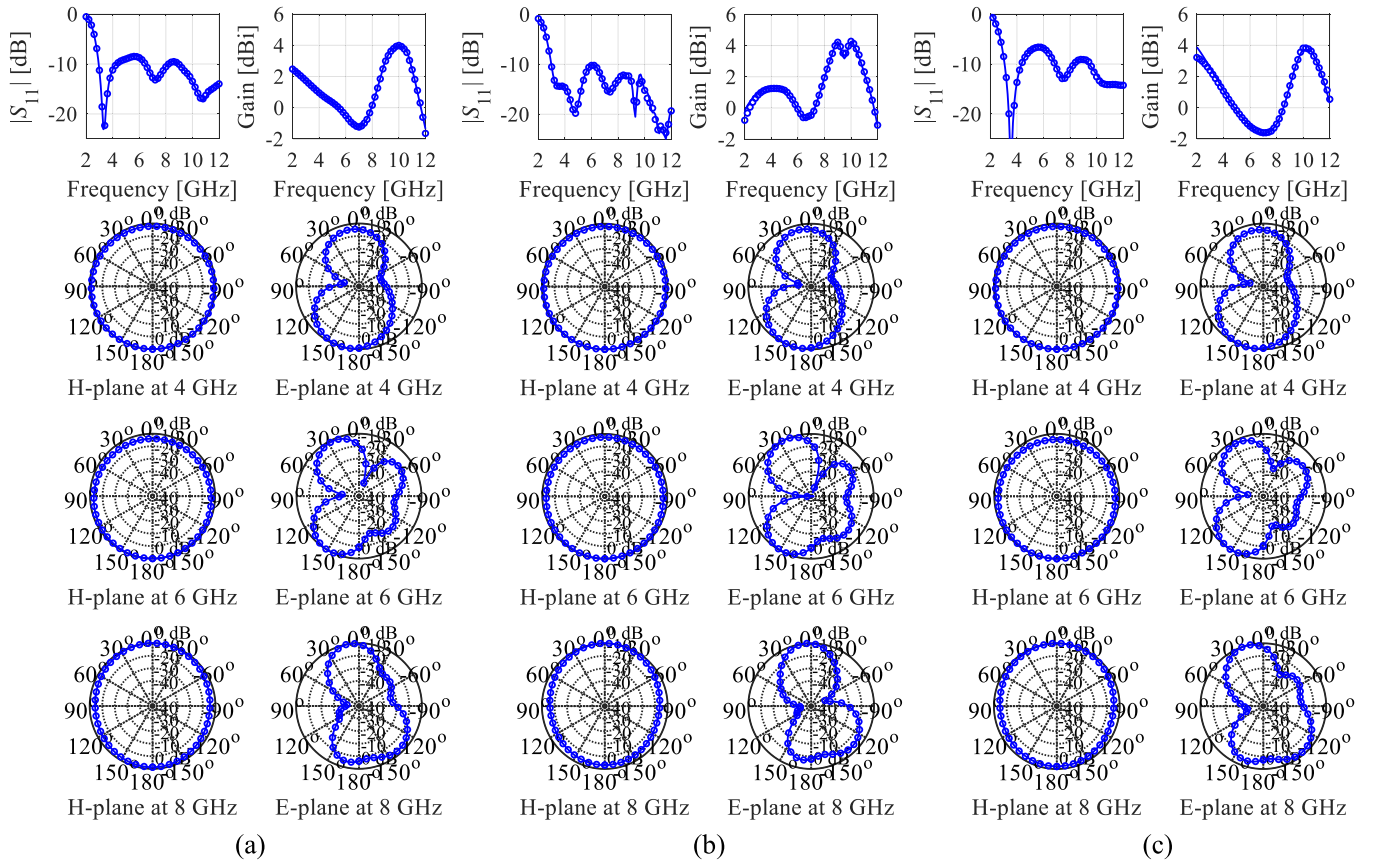


Fig. 9. Antenna I: Model-predicted (o) and EM-analysed antenna characteristics (—) at representative test points. Radiation patterns are shown at frequencies 4.0, 6.0, and 8.0 GHz. The model was built with 2000 training points: (a)–(c) design 1, 2, and 3.

allocated in the parameter space with the help of a modified Latin Hypercube Sampling (LHS) [148]. In Section 3, we will consider different numbers of training points to investigate the surrogate’s scalability. Given $\{\mathbf{x}_B^{(j)}\}$, we acquire the antenna responses $\{\mathbf{R}_{EM}(\mathbf{x}_B^{(j)}), \mathbf{x}_B^{(j)}\}$, $j = 1, \dots, N_B$. The specific datasets representing reflection, gain, and radiation patterns are $\{S_{11}(\mathbf{x}_B^{(j)}, f), \mathbf{x}_B^{(j)}\}$, $\{G(\mathbf{x}_B^{(j)}, f), \mathbf{x}_B^{(j)}\}$, and $\{P(\mathbf{x}_B^{(j)}, f, \theta), \mathbf{x}_B^{(j)}\}$. For the sake of brevity, the latter stands for both H- and E-plane patterns. The modeling process block diagram is visualized in Fig. 6.

3. Results

This part of the paper illustrates the presented modeling methodology using the two broadband antennas described in Section 3.1. The setup and benchmark approaches are detailed in Section 3.2. The numerical results are encapsulated in Section 3.3. The fundamental findings are presented in Section 3.4, while Section 3.5 focuses on the design applications of the developed model for antenna optimization under various objectives and performance criteria, such as improving impedance matching, minimizing antenna size, and achieving omnidirectional radiation. Finally, Section 3.6 reports on the prototyping and measurement of selected designs optimized using the presented metamodel.

3.1. Test antennas

Figs. 7 and 8 present the parameterized geometries of the test antennas, along with key design information, including design variables, substrate properties, and parameter ranges. Both antennas are compact ultra-wideband (UWB) monopoles featuring modified ground planes and radiator slots, designed to operate within the 3.1–10.6 GHz band. The responses of interest include the reflection coefficient, broadside realized gain, and far-field patterns in two principal planes: the H-plane

(perpendicular to the feeding line) and the E-plane (perpendicular to the substrate and aligned with the feed line). The frequency range analyzed spans 2.0–12.0 GHz, with azimuthal angles from -180° to $+180^\circ$. Electromagnetic simulations are executed using CST Microwave Studio. The computational models include SMA connectors.

3.2. Setup and Benchmark Methods

To assess the scalability of the presented modeling strategy, surrogates for Antennas I and II were built using three different dataset sizes: $N_B = 500, 1000$, and 2000 . The training samples were generated using LHS [148]. The predictive performance of the models was quantified using RRMSE, as defined in Section 2.1, and computed on an independent test set of 200 samples, also generated using LHS. The proposed technique is juxtaposed against several approaches listed in Table 4, including regression-based approaches such as SVR and GPR, as well as deep learning models, including MLPs, CNNs, and RNNs.

3.3. Results

The modeling outcomes are encapsulated in Tables 5 and 6. The data encompasses the RRMSE values obtained for radiation pattern modeling using the proposed technique and benchmark methods for the three training datasets (500, 1000, and 2000 samples). The errors are given for each type of radiation pattern considered (H-plane and E-plane). Furthermore, we also consider the mean values that represent the average performance of the respective surrogates.

The RRMSE obtained for reflection responses S_{11} and gain using the proposed method are 8.5% and 7.0% ($N_B = 500$), 7.2% and 5.5% ($N_B = 1000$), and 5.2% and 3.4% ($N_B = 2000$) for Antenna I and similar for Antenna II. These figures, especially for the largest training set, are more

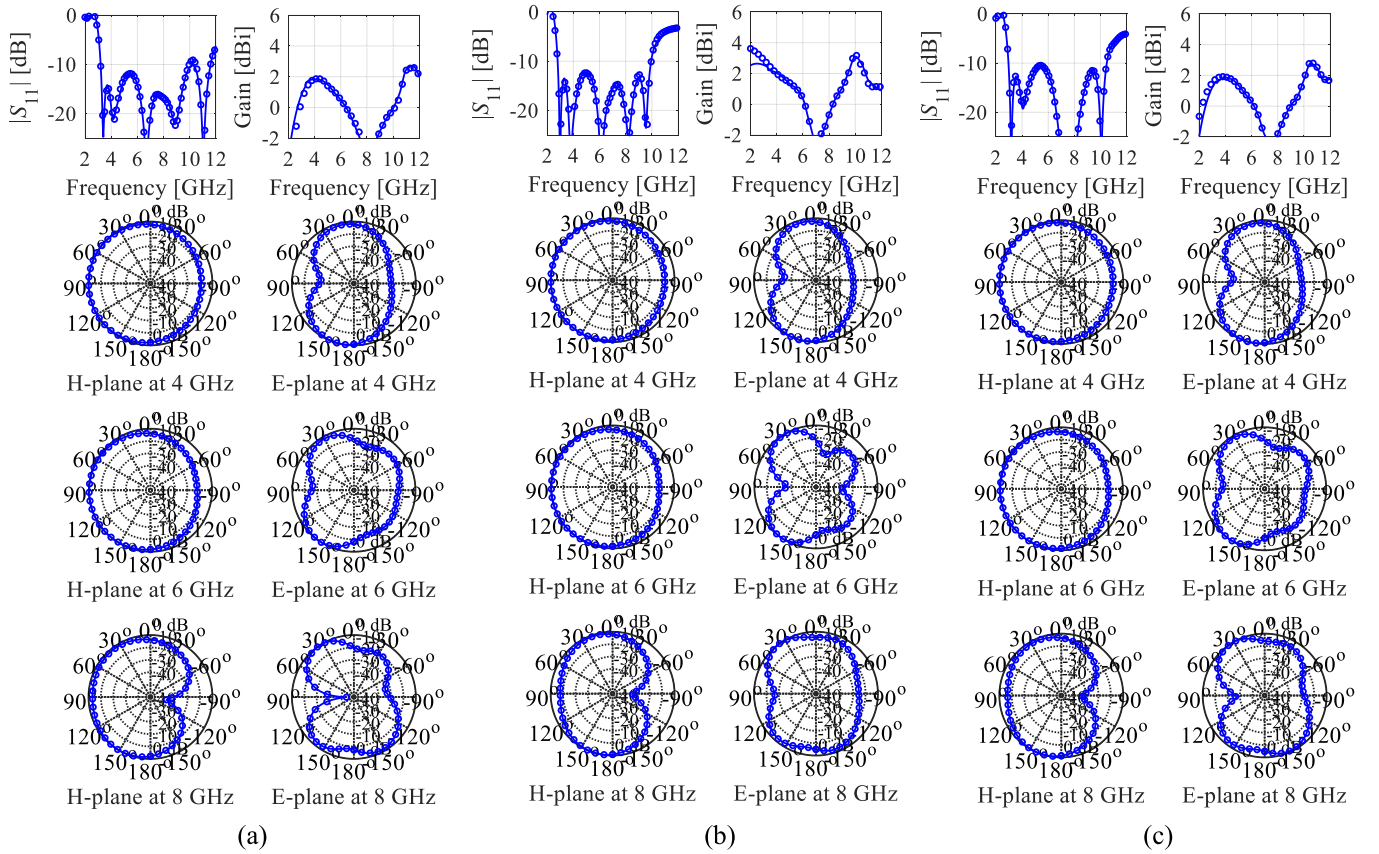


Fig. 10. Antenna II: Model-predicted (o) and EM-analysed antenna characteristics (—) at representative test points. Radiation patterns are shown at frequencies 4.0, 6.0, and 8.0 GHz. The model was built with 2000 training points: (a)–(c) design 1, 2, and 3.

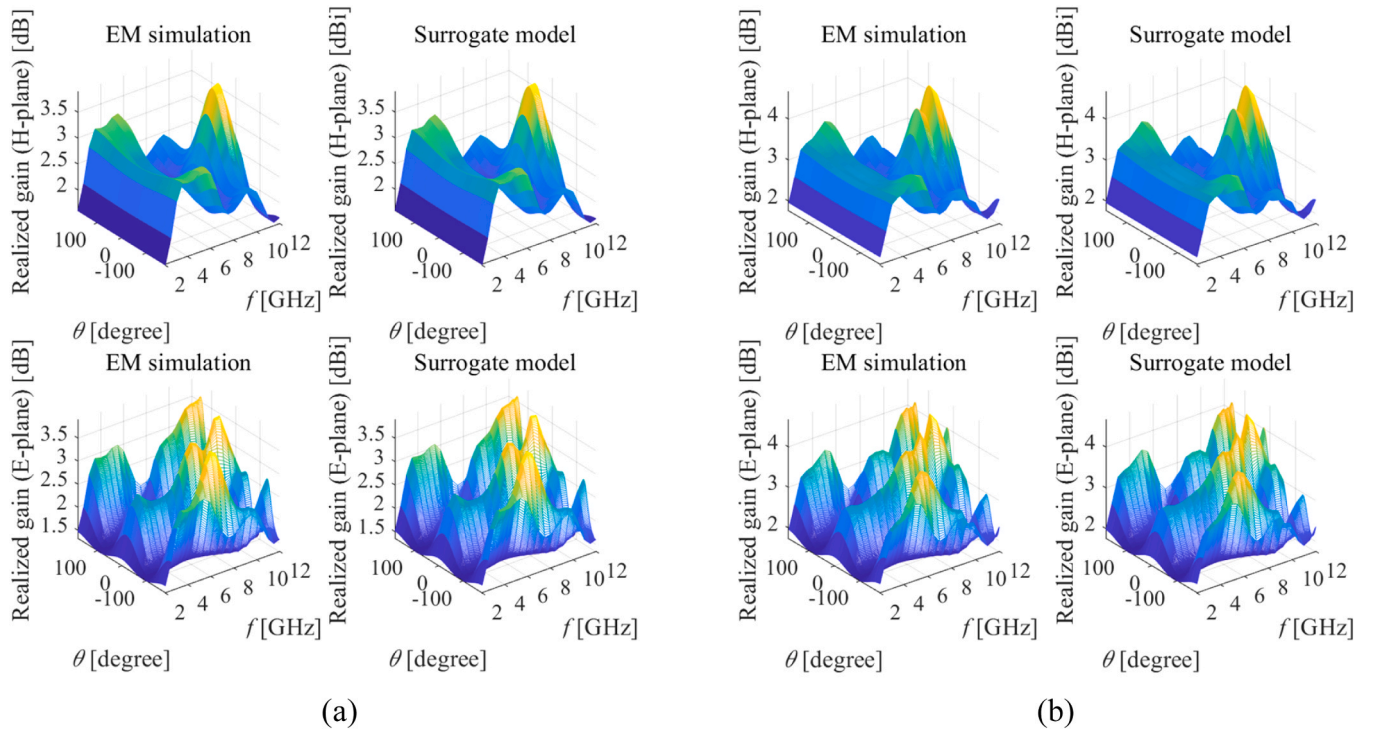


Fig. 11. Antenna I: radiation pattern surfaces as functions of frequency f and azimuth angle θ for selected test locations: EM simulation versus presented model constructed using 2000 samples: (a) H-plane, (b) E-plane: (a) design 1, (b) design 2.

than sufficient to ensure suitability of the model for design applications.

Figs. 9 and 10 visualize the responses of Antennas I and II for selected

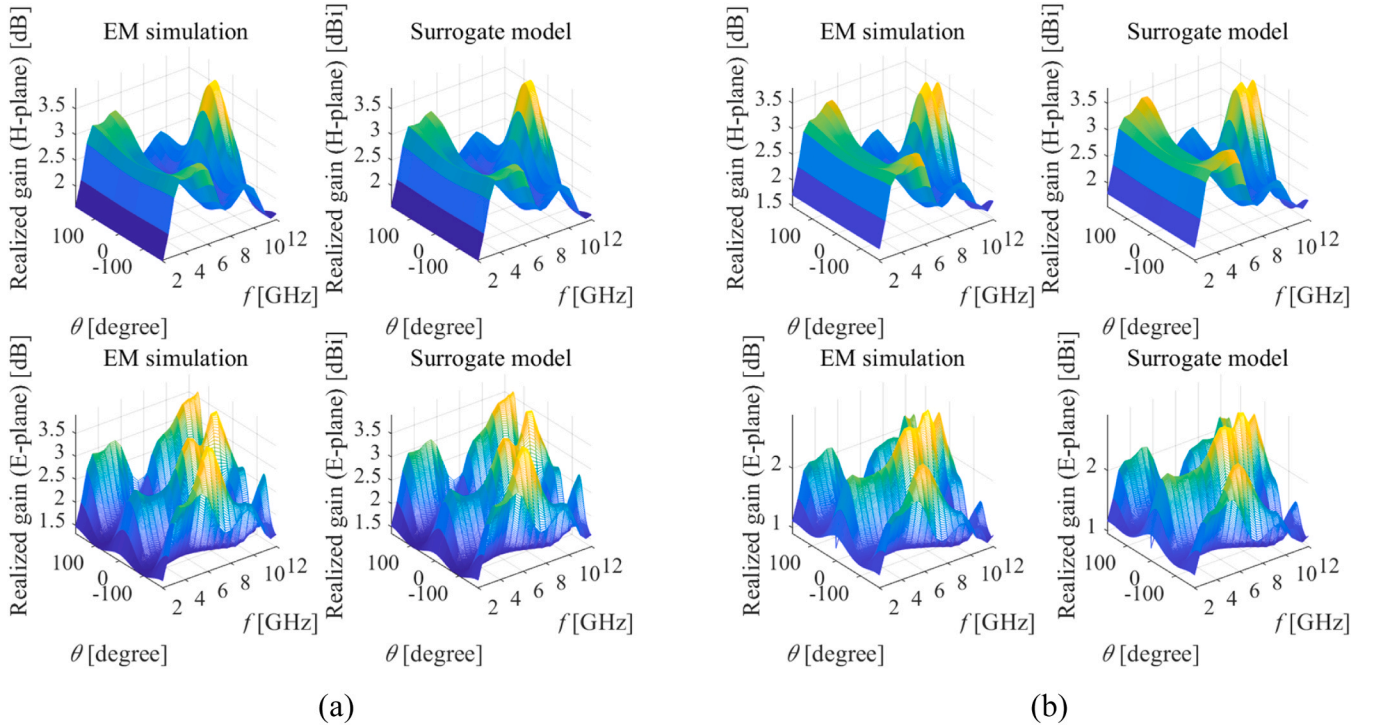


Fig. 12. Antenna II: radiation pattern surfaces as functions of frequency f and azimuth angle θ for selected test locations: EM simulation versus presented model constructed using 2000 samples: (a) H-plane, (b) E-plane: (a) design 1, (b) design 2.

Table 7

S_{11} and Gain Modeling Accuracy Using Various NN Models (Antenna I, $N_B = 2000$).

Model	S_{11}	Gain
ANN	17.3%	14.7%
RNN-I	5.6%	7.0%
RNN-II	10.2%	9.1%
Transformer	11.9%	9.5%
CNN	5.2%	3.4%

Table 8

Training and Inference Time for Antenna 1 (Model Built Using $N_B = 500, 1000$, and 2000 Data Samples).

Modeling method	Training time (in hours) for $N_B = 500$	Training time (in hours) for $N_B = 1000$	Training time (in hours) for $N_B = 2000$	Inference/Prediction time (per sample)
LR	0.01	0.03	0.1	< 1 s
GPR	0.3	0.8	2	
SVR	0.01	0.07	0.2	
FF-NN	0.2	0.7	1.5	
RNN_1	0.4	1.1	2.5	
RNN_2	0.6	1.4	3.75	
CNN	0.8	1.2	3.5	
Transformers	1.0	2.8	5	
This work	2.5	3.75	8	

testing samples. The surrogate-predicted and EM-simulated radiation pattern surfaces, parameterized by frequency f and the azimuth angles θ , can be found in Figs. 11 and 12.

A remark should be made regarding the modelling of S_{11} and gain. The selection of CNN model for this purpose was based on both architectural considerations and benchmark performance. The convolutional structure of CNN enables efficient extraction of local features from the frequency responses while providing a regularized mapping suitable for

Table 9

Performance figures for test antennas.

Performance figure	Symbol	Unit	Explanation
Impedance matching	S_M	dB	Maximum in-band reflection coefficient (here, over 3.1 GHz to 10.6 GHz band)
Antenna size	A	mm ²	Footprint area, i.e., the size of antenna substrate
Gain variability	ΔG	dB	Difference between maximum and minimum broadside gain over the operating bandwidth (here, from 3.1 GHz to 10.6 GHz)
Omnidirectionality	ΔR	dB	Difference between maximum and minimum H-plane farfield radiation for three selected frequencies of 4 GHz, 6 GHz, and 8 GHz

limited training data (see also Section 2.2.3). As shown in Table 7, the CNN achieves the lowest prediction error for both S_{11} and gain among all considered deep learning models, including MLP model, two types of RNNs, and the Transformer model. Thus, for the problem considered in this work, the CNN model provides the best accuracy–complexity trade-off.

The computational cost of the training of both benchmark models and the proposed one was also assessed (see Table 8). Approximate timings were obtained on the following hardware platform: Intel Xeon CPU, 2.1 GHz, 64 GB RAM, without GPU acceleration. The reported times correspond to Antenna I and are provided for all three training dataset sizes, $N_B = 500, 1000$, and 2000 samples to illustrate the increase in computational cost with training set size. As expected, the most offline cost was incurred for the current work, due to its complex multi-branch recurrent, attention, and convolutional-decoder architecture. Nevertheless, for all models, the average prediction time remains below one second per test sample, which is particularly important for surrogate-assisted design involving numerous model evaluations.

Table 10
Optimization of Antenna I.

Case	Design specifications			Results (performance figures)			
	Primary objective	Constraint 1	Constraint 2	S_M [dB]	A [mm ²]	ΔG [dB]	ΔR [dB]
1	Impedance matching improvement	-	-	-15.2	454	3.1	6.4
2	Size reduction	$S_M \leq -10$ dB	-	-9.9	171	4.8	6.6
3	Size reduction	$S_M \leq -12$ dB	-	-11.8	226	5.2	5.8
4	Size reduction	$S_M \leq -14$ dB	-	-13.6	296	5.5	6.4
5	Gain variability reduction	$S_M \leq -10$ dB	-	-10.0	398	2.4	7.2
6	Gain variability reduction	$S_M \leq -10$ dB	$A \leq 250$ mm ²	-10.0	243	3.9	8.7
7	Omnidirectional pattern enhancement	$S_M \leq -10$ dB	-	-10.0	359	4.1	3.3
8	Impedance matching improvement	$\Delta R \leq 4$ dB	-	-11.2	344	4.2	4.0

Table 11
Optimization of Antenna II.

Case	Design specifications			Results (performance figures)			
	Primary objective	Constraint 1	Constraint 2	S_M [dB]	A [mm ²]	ΔG [dB]	ΔR [dB]
1	Impedance matching improvement	-	-	-13.6	453	6.2	33.6
2	Size reduction	$S_M \leq -10$ dB	-	-10.0	309	5.2	19.9
3	Size reduction	$S_M \leq -12$ dB	-	-11.9	385	4.8	36.7
4	Size reduction	$S_M \leq -14$ dB	-	-12.9	403	5.7	34.2
5	Gain variability reduction	$S_M \leq -10$ dB	-	-10.0	409	4.1	33.4
6	Gain variability reduction	$S_M \leq -10$ dB	$A \leq 350$ mm ²	-10.0	350	4.8	28.8
7	Omnidirectional pattern enhancement	$S_M \leq -10$ dB	-	-10.0	310	5.2	19.8
8	Impedance matching improvement	$\Delta R \leq 25$ dB	-	-11.6	337	5.2	25.0

3.4. Discussion

The results underscore the competitive performance of the presented modeling approach. On the one hand, it achieves lower RRMSE values than all benchmark methods. On the other hand, its superiority remains consistent across both test antennas and all considered training dataset sizes.

The obtained results also demonstrate a clear reduction in modeling error as the number of training samples increases. The error levels achieved, particularly for the largest training datasets of 2000 samples, are remarkably low (below 1% for Antenna I and below 2% for Antenna II). This accuracy level makes the proposed surrogate suitable for handling antenna design tasks, as discussed in detail in Section 3.5. Importantly, satisfactory performance is also maintained for models constructed using smaller training datasets (500 and 1000 samples), which confirms the robustness and data efficiency of the proposed approach.

The proposed model was further evaluated against several benchmark methods to demonstrate its predictive advantage over established regression and neural-network-based surrogate models. Beyond providing direct performance comparisons, the selected benchmarks were designed to represent the key architectural mechanisms incorporated into the proposed AFN. Specifically, the RNN-based baselines assess recurrent processing without the proposed cross-attention fusion and convolutional decoder. While the CNN baseline evaluates convolutional response reconstruction without recurrent attention-based feature extraction. Finally, the Transformer baseline examines attention-based modeling without the proposed multi-branch recurrent representation. The superior performance of the proposed AFN therefore confirms the effectiveness of integrating all complementary components within a unified surrogate modeling framework.

The benchmark methodologies, including traditional regression models and deep learning ones, are significantly inferior. For instance, the deep learning metamodel involving transformers, the best benchmark technique for Antenna I, exhibits an RRMSE that is almost three times higher than that of the proposed approach for $N_B = 2000$. Similarly, for Antenna II, the best benchmark procedures, SVR and GPR, exhibit an RRMSE about twice that of our method (also for $N_B = 2000$). Other benchmark techniques are even more inferior. The advantages of

our methodology are even more pronounced for $N_B = 500$, especially for Antenna II, where its RRMSE is typically four to six times lower than the benchmark. These results underscore the relevance of the algorithmic mechanisms used in the presented surrogate, as elaborated in Section 2.

3.5. Design applications

The most important aspect of surrogate modeling, often overlooked in literature, is practical (design) applications. This section of the article demonstrates the application of our metamodels for optimizing test antennas under various scenarios. The requirements (and constraints) are imposed upon the reflection responses, physical dimensions, broadside gain, and radiation patterns. Table 9 provides definitions of performance figures, specifically impedance matching, antenna size, gain variability, and pattern omnidirectionality. Tables 10 and 11 present the design cases considered. In each scenario, a primary objective is selected, and it may be optionally accompanied by constraints. The default operating frequency range for both antennas is UWB. The solutions are obtained by optimizing the presented surrogate (an $N_B = 2000$ version thereof). The respective designs are subject to EM analysis. The EM-evaluated performance figures are reported in the right-hand-side columns of Tables 10 and 11, whereas Figs. 13 and 14 provide a comparison between model-predicted and EM-evaluated responses. Here, a particle swarm optimizer (PSO) [157] is employed to conduct a global search, and a gradient-based routine [158] is used for final tuning. It should be emphasized that other bio-inspired routines can be applied and run with a large computational budget because the search is carried out using predictions of a fast surrogate (i.e., its optimization cost is negligible).

The results, as encapsulated in Tables 10 and 11, indicate the versatility of the proposed surrogate model. As shown, the test antennas can be optimized for various design specifications, including impedance matching enhancement, size reduction, gain variability reduction, and achieving omnidirectional radiation patterns. The optimization process may be subject to auxiliary constraints, typically related to reflection coefficient $|S_{11}| \leq -10$ dB (the primary requirement), but also other conditions such as maximum size, etc. The alignment between the metamodel model predictions and EM analysis is excellent for all scenarios, indicating that the designs produced by optimizing the

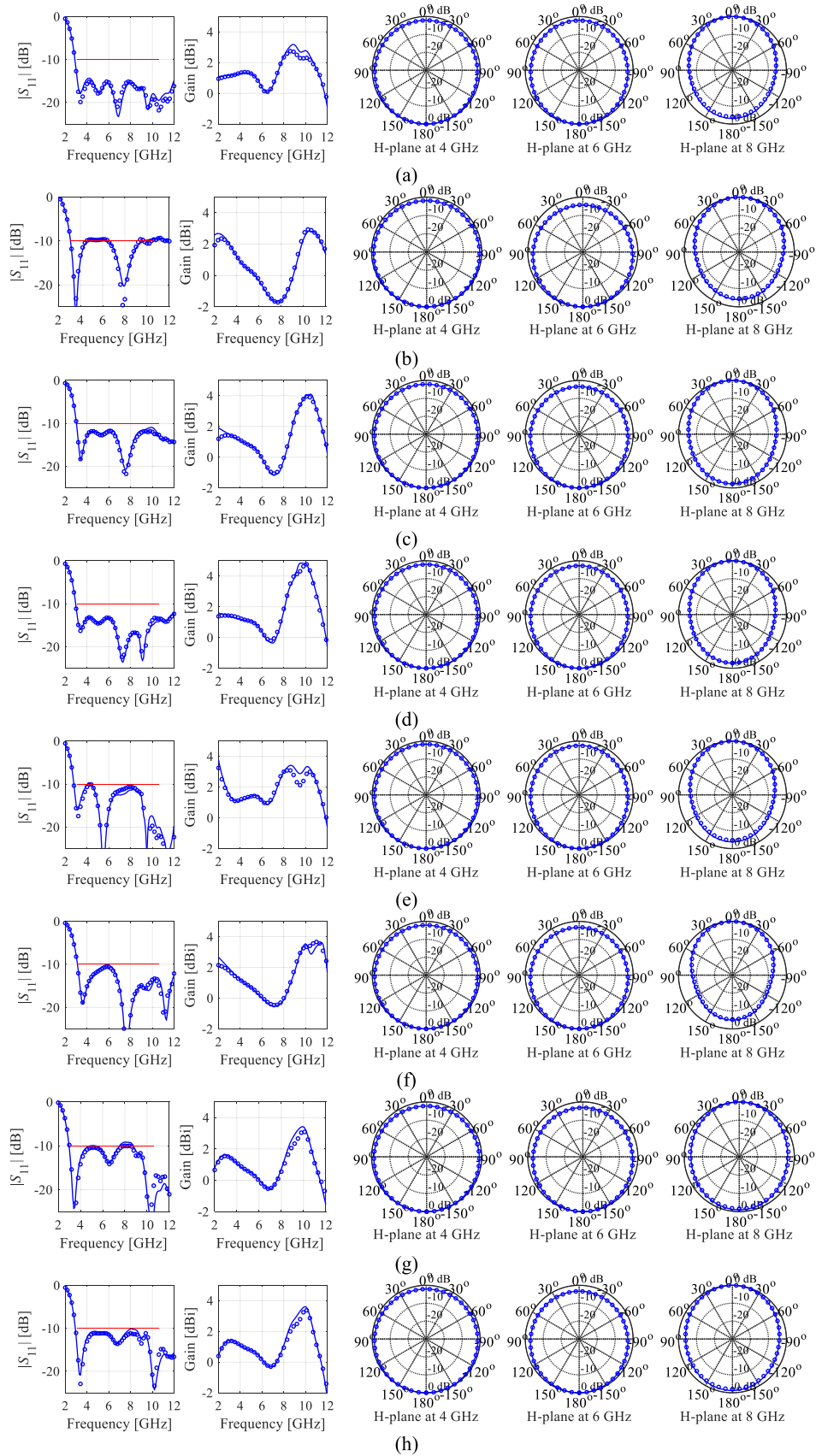


Fig. 13. Optimization of Antenna I: (a)–(h) Case 1 through 8. The pictures show surrogate-predicted (o) and EM-evaluated (—) outputs at the solutions generated by optimizing the proposed metamodel ($N_B = 2000$). Only H-plane radiation patterns are shown.

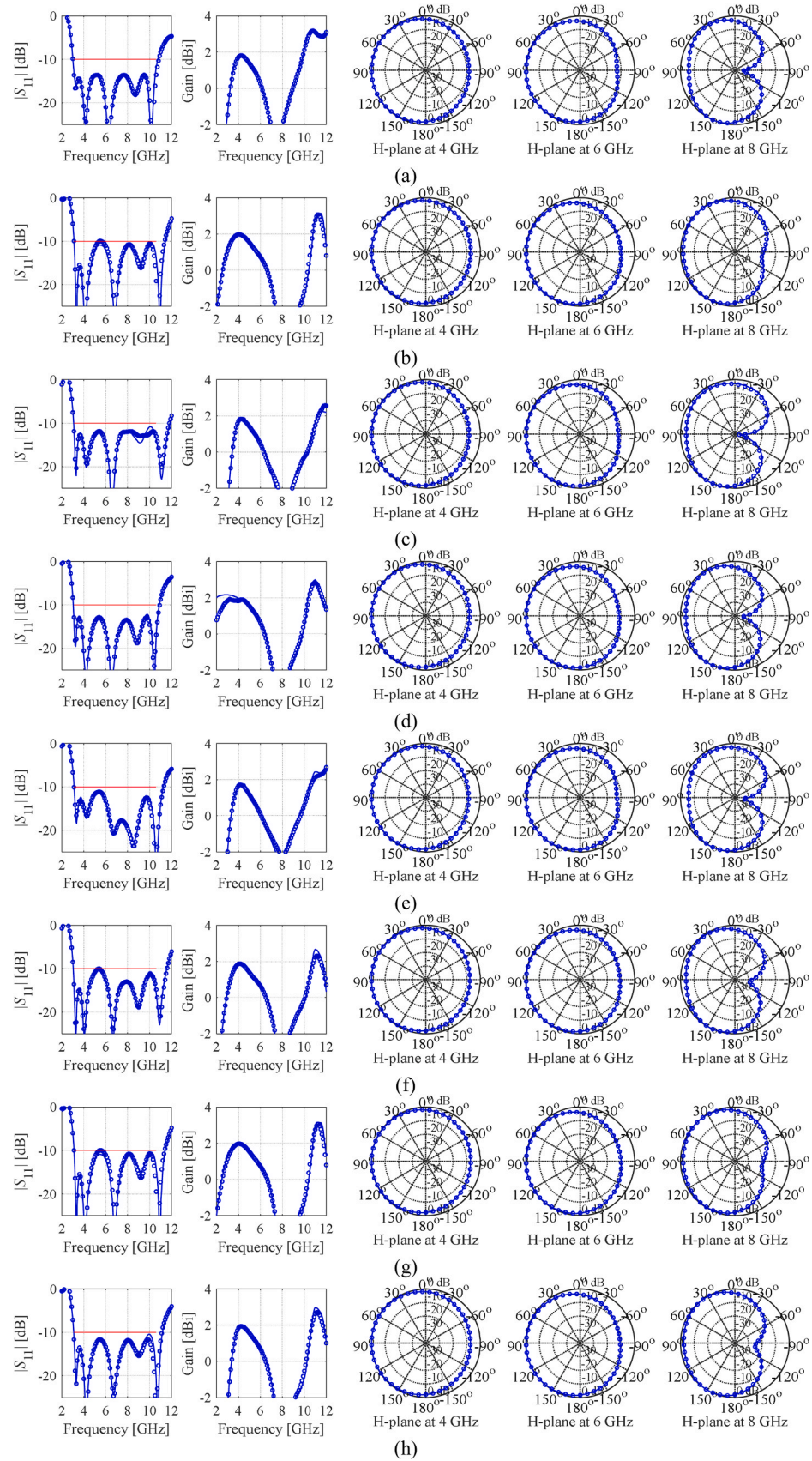


Fig. 14. Optimization of Antenna II: (a)–(h) Case 1 through 8. The pictures show surrogate-predicted (o) and EM-evaluated (—) outputs at the solutions generated by optimizing the proposed metamodel ($N_B = 2000$). Only H-plane radiation patterns are shown.

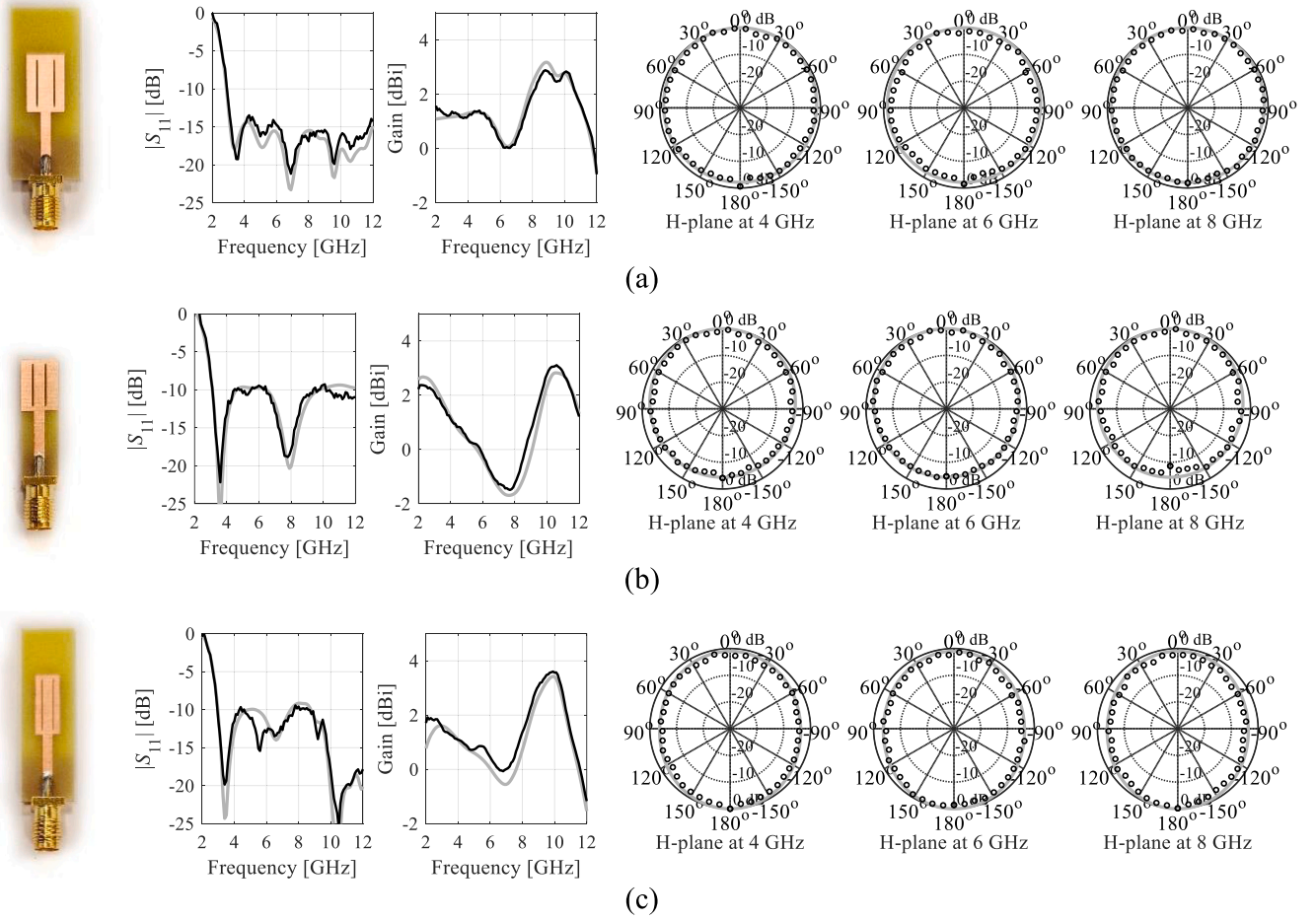


Fig. 15. Experimental validation of Antenna I: (a) Case 1, (b) Case 2, (c) Case 7. EM simulated and measured characteristics are distinguished using grey and black colors.

metamodel do not require further tuning.

As observed, a topologically identical antenna structure can be designed for diverse requirements. For example, the best in-band $|S_{11}|$ for Antenna I is as good as -15 dB (Case 1) with the corresponding structure size of 454 mm^2 . On the other hand, the size may be reduced to 171 mm^2 while still preserving -10 dB UWB bandwidth (Case 2). Specific performance figures may be improved (at the expense of other parameters). For example, the broadside gain variability may be reduced to approximately 2 dB while enlarging the size and sacrificing some omnidirectionality (Case 5). Alternatively, the best omnidirectionality (with ΔR just 3 dB) is obtained at larger gain variation (Case 7). For Antenna II, similar design trade-offs can be observed, although the L-shaped ground plane stub prevents achieving a perfectly omnidirectional pattern at 8 GHz. Still, the radiation null at $\theta = -90^\circ$ can be significantly reduced (Case 7) compared to the best impedance matching design (Case 1). For this structure, it appears that the best omnidirectionality corresponds to the minimum size design (Case 2).

It should also be emphasized that setting up appropriate constraints effectively enables rapid multi-objective optimization. For example, Cases 1 through 4 for both test antennas represent Pareto sets with respect to two goals: F_1 – minimization of the maximum $|S_{11}|$, and F_2 – size reduction. Similar trade-offs can be generated by other arrangements of constraints and design goals. Note that all these results are produced by reusing the same surrogate model without incurring any extra computational costs. The key factors are extensive coverage of the parameter space (allowing for the accommodation of optimum designs regarding different performance figures) and excellent surrogate

accuracy.

3.6. Experimental Validation

For supplementary illustration, selected designs of Antennas I and II were fabricated and experimentally validated: (a) Case 1 (best impedance matching), (ii) Case 2 (minimum size), and (iii) Case 7 (best omnidirectionality). The antenna prototypes are illustrated in Fig. 15 (Antenna I) and Fig. 16 (Antenna II). The figures also provide a comparison between EM-evaluated and measured $|S_{11}|$, gain, and H-plane patterns. As observed, the dataset alignment is good. Small discrepancies are attributable to manufacturing and assembly inaccuracies as well as measurement errors (cable effect, imperfect VNA calibration, etc.). To improve the reliability of EM analysis, the computational models are equipped with SMA connectors, which are particularly important for electrically small antennas, where the bulk of connector's metal may significantly affect antenna characteristics, especially impedance matching.

4. Conclusion

The purpose of this study was to develop a deep-learning-based procedure for comprehensive behavioral modeling of broadband antennas. The primary objective was to accurately predict electrical responses and far-field radiation patterns across a wide frequency range, a broad angular span, and an extensive domain of geometry parameters. Once trained, the resulting surrogates exhibit high versatility and can be

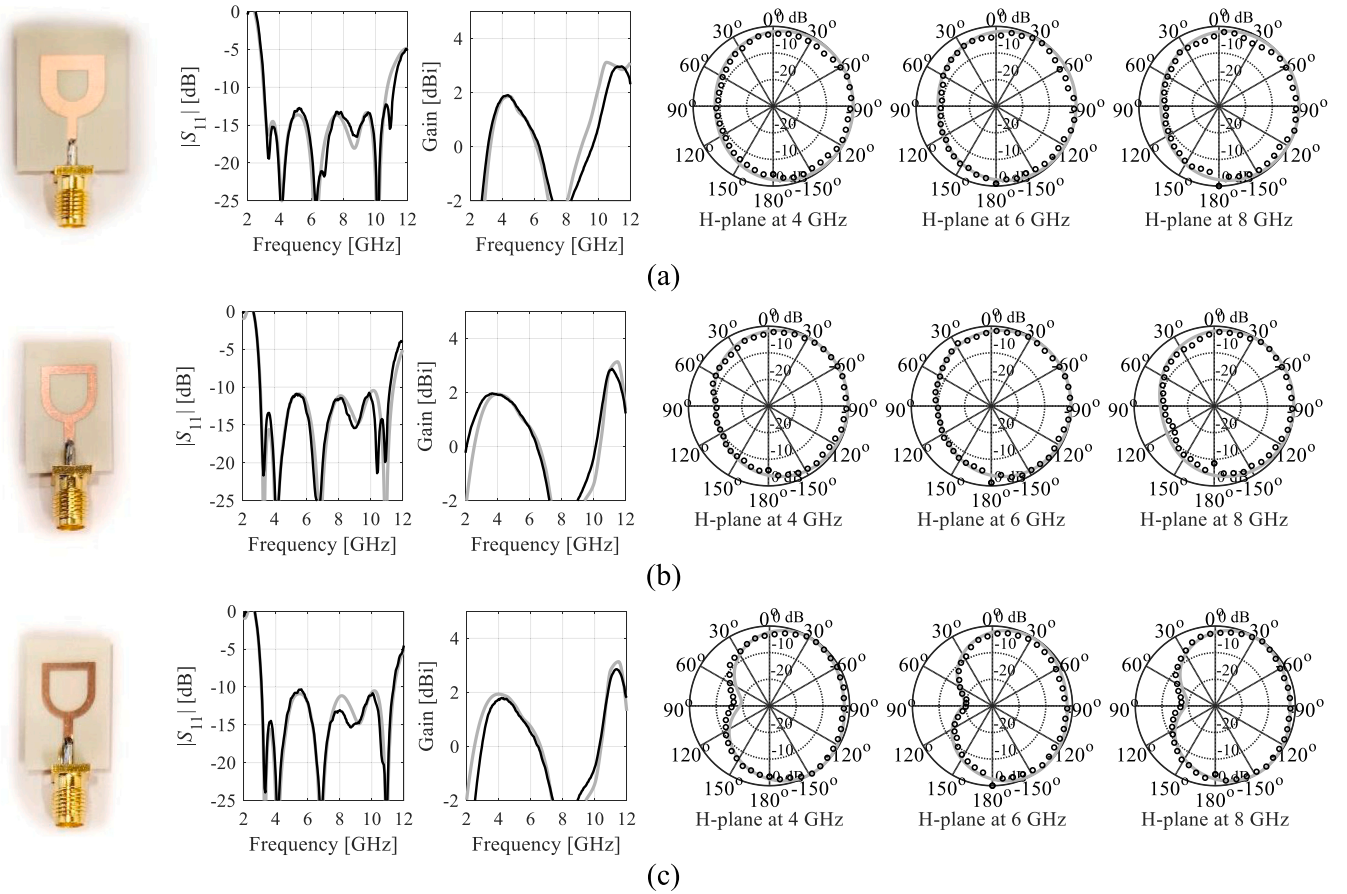


Fig. 16. Experimental validation of Antenna II: (a) Case 1, (b) Case 2, (c) Case 7. EM simulated and measured characteristics are distinguished using grey and black colors.

readily employed to solve a wide array of antenna design and optimization problems. Reflection and gain responses are modeled using a conventional convolutional neural network (CNN); however, radiation pattern prediction is handled by a custom hybrid architecture that integrates multi-branch recurrent feature extraction with cross-attention-driven fusion. Specifically, six independent RNN branches, each combining fully connected, LSTM, and GRU layers, capture complementary geometric dependencies. Their embeddings are merged via a cascade of attention modules and subsequently decoded with a transposed convolutional refinement network, which can reconstruct complete H- and E-plane patterns.

Our technique was validated using two UWB antennas evaluated over the 2–12 GHz band and across broad ranges of geometric parameters. Extensive numerical studies, including training sets of 500, 1000, and 2000 samples and comparisons with numerous benchmarks (regression models and CNN-, RNN-, and Transformer-based frameworks), demonstrate exceptional predictive accuracy. The achieved RRMSE is between 1% and 2%, whereas the most competitive benchmark methods exhibit errors at least three times larger, with average performance lagging even further behind. To illustrate the practical utility, the surrogates were used to optimize the test antennas under diverse scenarios, including improving impedance matching, reducing footprint, stabilizing gain, enhancing omnidirectionality, and performing multi-objective design (e.g., exploring trade-offs between in-band reflection and physical size). The results confirm the robustness, flexibility, and design relevance of the presented models.

Overall, the proposed modeling strategy represents a substantial advancement in the full-spectrum behavioral modeling of antenna systems, supporting the broader integration of data-driven tools into

modern antenna design processes. Despite the demonstrated effectiveness of the proposed methodology, certain limitations exist that define directions for future research. As with other data-driven neural-network surrogates, the validity of the proposed model is tied to the antenna topologies, parameter ranges, and response representations covered by the training data. If the design-space dimensionality, parameter ranges, or output response dimensionality are substantially increased, larger EM-simulated datasets would be required to maintain accuracy and generalization. Within the considered domain, the versatility of the approach has been demonstrated using two broadband antenna structures and multiple design tasks, including impedance matching, size reduction, gain stabilization, and omnidirectional-pattern improvement. However, extension beyond the present modeling domain would require additional training samples and retraining. Although the proposed modeling approach has been applied to broadband antennas, in principle, it can handle other types of antenna systems, in particular, directive antennas such as quasi-Yagi structures, dielectric resonator antennas (DRAs), or antenna arrays. However, handling more complex scenarios, especially those involving higher-dimensional parameter spaces, will require certain modifications, primarily, appropriate dimensionality reduction approaches and/or the employment of regularization techniques such as the response feature technology. Addressing these aspects will be the focus of future work.

Thus, future studies will focus on extending the method's scalability to higher-dimensional problems and reducing training costs while maintaining accuracy and reliability. Additionally, future work will focus on extending the method to full 3D radiation-pattern reconstruction using frequency–azimuth–elevation response tensors, and on reducing training costs while maintaining accuracy and reliability.

Furthermore, efficient attention variants for Transformers may also be explored for higher-dimensional response representations or longer latent sequences. Quantum-inspired neural computing may also be considered in future work with respect to scalability and computational complexity [159]. The extension of the presented modeling approach to other antenna types, especially directive antennas, will also be considered in future work. Improved modeling procedure will also allow to consider more complex real-world optimization scenarios that take into account multiple constraints (e.g., acceptance thresholds for selected antenna responses such as gain, axial ratio, reflection coefficients, etc.).

CRedit authorship contribution statement

Slawomir Koziel: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kaustab C. Sahu:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Anna Pietrenko-Dabrowska:** Writing – review & editing, Supervision, Software, Resources, Project administration, Funding acquisition, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgment

The authors thank Dassault Systemes, France, for making CST Microwave Studio available. This work was partly supported by the Icelandic Research Fund Grant 2410297 and by the National Science Centre of Poland Grant 2025/57/B/ST7/00738.

Data availability

No data was used for the research described in the article.

References

- [1] M. Agrawal, R. Yadav, S. Koziel, High gain and high isolation SIW-based dual-circularly polarized metasurface antenna using CMA-based mode control for MMW applications, *IEEE Trans. Ant. Propag.* (2025).
- [2] S.-H. Cheng, S.-C. Chen, W.-Y. Huang, Low-profile MIMO trapezoidal patch antenna for 5G wideband mobile antenna application, *IEEE Ant. Wirel. Propag. Lett.* 24 (3) (2025) 696–700.
- [3] B.A.F. Esmail, S. Koziel, Design and optimization of metamaterial-based highly-isolated MIMO antenna with high gain and beam tilting ability for 5G millimeter wave applications (paper no), *Sc. Rep.* 14 (2024) 3203.
- [4] A. Hossain, S. Pancrazio, T. Kelley, A.-V. Pham, A compact and low-profile high-gain multilayer Vivaldi antenna based on gradient metasurface superstrates, *IEEE Ant. Wirel. Propag. Lett.* 24 (6) (2025) 1537–1541.
- [5] Z. Wu, L. Li, W. Wu, J.J. Liang, L.L. Huang, Ultrawideband polarization-reconfigurable antenna based on C-shaped radiators, *IEEE Ant. Wirel. Propag. Lett.* 24 (1) (2025) 113–116.
- [6] K.-D. Jang, J.-H. Lee, Frequency-tunable tri-band metamaterial-inspired antenna, *IEEE Access.* 13 (2025) 21210–21215.
- [7] I. Khan, K. Zhang, L. Ali, Q. Wu, Enhanced quad-port MIMO antenna isolation with metamaterial superstrate, *IEEE Ant. Wirel. Propag. Lett.* 23 (1) (2024) 439–443.
- [8] J.-M. Poyanco, A. Fernández-Durán, E. Rajo-Iglesias, Unified design: Sub-6 GHz conformal patch antennas integrated with millimeter-wave reflector for fixed wireless access terminals, *IEEE Trns. Ant. Prop.* 73 (1) (2025) 635–640.
- [9] Y. Wang, S.W. Qu, Low-profile dual-circularly polarized antenna array with hemispherical scan coverage, *IEEE Ant. Wirel. Propag. Lett.* 23 (4) (2024) 1191–1195.
- [10] J.D. Peng, Y. Li, L.H. Ye, J.F. Li, D.L. Wu, Shared dipoles inspired horizontally polarized omnidirectional antenna with broad bandwidth, *IEEE Ant. Wirel. Propag. Lett.* 23 (8) (2024) 2386–2390.
- [11] T.Y. Yan, X.H. Ding, J.Y. Yang, J.X. Chen, A low-cost compact dual-polarized patch antenna array for 5G massive MIMO base station, *IEEE Ant. Wirel. Propag. Lett.* 23 (4) (2024) 1381–1385.
- [12] S. Song, X. Chen, A compact and broadband dielectric resonator magnetoelectric dipole antenna with high front-to-back ratio, *IEEE Ant. Wirel. Propag. Lett.* 23 (1) (2024) 239–243.
- [13] F.-F. Fan, Q.-L. Chen, Y.-X. Xu, X.-F. Zhao, J.-C. Feng, Z.-H. Yan, A wideband compact printed dipole antenna array with SICL feeding network for 5G application, *IEEE Ant. Wirel. Propag. Lett.* 22 (2) (2023) 283–287.
- [14] V. Jayaprakash, C. Ds, R.K. Barik, S. Koziel, Compact substrate-integrated hexagonal cavity-backed self-hexaplexing antenna for Sub-6 GHz applications, *IEEE Access.* 12 (2024) 54414–54421.
- [15] B.A.P. Nel, A.K. Skriverik, M. Gustafsson, Radiation efficiency and gain bounds for microstrip patch antennas, *IEEE Trans. Ant. Prop.* 73 (2) (2025) 873–883.
- [16] H. Koo, S. Nam, Ku-band printed bow-tie antenna H-plane array with low mutual coupling, *IEEE Trans. Ant. Propag.* 72 (1) (2024) 980–985.
- [17] S. Koziel, S. Ogurtsov, Simulation-based optimization of antenna arrays, World Scientific, Singapore, 2019.
- [18] M.A. Haq, S. Koziel, Feedline alterations for optimization-based design of compact super wideband MIMO antennas in parallel configuration, *IEEE Ant. Wirel. Prop. Lett.* 18 (10) (2019) 1986–1990.
- [19] Y. Zhang, N. Liu, X. Sheng, Z. Zhao, H. Ding, An efficient curved patches-based meshing method for calculating EM performance of BoR radomes, *IEEE Ant. Wirel. Propag. Lett.* 24 (2) (2025) 319–323.
- [20] M.W. Wolff, J.A. Nanzer, Application of pseudoweights in antenna array optimization and design, *IEEE Ant. Wirel. Propag. Lett.* 23 (5) (2024) 1478–1482.
- [21] L.-C. Chang, Y.-H. Lin, S.-Y. Chen, H.-T. Chou, H. Wang, A duplexing hybrid slot antenna design with high isolation for short-range radar detection and identification applications at 24 GHz band, *IEEE Trans. Ant. Propag.* 70 (4) (2022) 2468–2479.
- [22] U. Ullah, M. Al-Hasan, S. Koziel, I. Ben Mabrouk, Circular polarization diversity implementation for correlation reduction in wideband low-cost multiple-input-multiple-output antenna, *IEEE Access.* 8 (1) (2020) 95585–95593.
- [23] Y. Sun, et al., Data-driven Bayesian optimization framework for rapidly developing novel wideband, low-profile dipole antenna with 3-D-printed technology, *IEEE Trans. Ant. Propag.* 73 (1) (2025) 108–120.
- [24] X. Liu, K. Ning, S. Xue, L. Ge, K.W. Leung, J.F. Mao, Printed filtering dipole antenna with compact size and high selectivity, *IEEE Trans. Ant. Propag.* 72 (3) (2024) 2355–2367.
- [25] W. Lu, et al., Design of microwave ablation antenna for flexible omnidirectional/directional ablation zone control, *IEEE Ant. Wirel. Prop. Lett.* 24 (1) (2025) 18–22.
- [26] K. Fu, K.W. Leung, Machine-learning-assisted swarm intelligence algorithm for antenna optimization with mixed continuous and binary variables, *IEEE Trans. Ant. Propag.* 73 (3) (2025) 1662–1673.
- [27] T.-Y. Gao, Y.-C. Jiao, Y.-X. Zhang, L. Zhang, A hybrid self-adaptive differential evolution algorithm with simplified bayesian local optimizer for efficient design of antennas, *IEEE Trans. Ant. Prop.* 73 (1) (2025) 391–404.
- [28] X. Teng, Y. Hou, C. Guo, S. Wang, K. Ma, Broadband low-profile multifunction dual-polarized dipole antenna and array for 5G and WiFi application, *IEEE Ant. Wirel. Propag. Lett.* 23 (6) (2024) 1849–1853.
- [29] A. Majumder, S. Chatterjee, S. Chatterjee, S. Sinha Chaudhari, D.R. Poddar, Optimization of small-signal model of GaN HEMT by using evolutionary algorithms, *IEEE Microw. Wirel. Comp. Lett.* 27 (4) (2017) 362–364.
- [30] O.N. Oyelade, A.E.-S. Ezugwu, T.I.A. Mohamed, L. Abualigah, Ebola optimization search algorithm: a new nature-inspired metaheuristic optimization algorithm, *IEEE Access.* 10 (2022) 16150–16177.
- [31] H. Zhao, Y.L. Ban, Y. Liu, L. Xu, J. Hu, A novel tiling technique via an evolutionary strategy for large-scale irregular arrays, *IEEE Trans. Ant. Propag.* 72 (4) (2024) 3175–3189.
- [32] J.E. Rayas-Sanchez, S. Koziel, J.W. Bandler, Advanced RF and microwave design optimization: a journey and a vision of future trends, *IEEE J. Microw.* 1 (1) (2021) 481–493.
- [33] S. Koziel, A. Pietrenko-Dabrowska, “Recent advances in accelerated multi-objective design of high-frequency structures using knowledge-based constrained modeling approach,” (paper no), *Knowl. -Based Syst.* 214 (106726) (2021).
- [34] W.T. Li, H.Q. Li, L. Li, Y.Q. Hei, X.W. Shi, Efficient machine learning-assisted multiobjective antenna design for internet-of-things applications, *IEEE Ant. Wirel. Propag. Lett.* 24 (5) (2025) 1263–1267.
- [35] Y. Zeng, X. Qing, M.Y.W. Chia, A wideband circularly polarized antenna with a nonuniform metasurface designed via multiobjective Bayesian optimization, *IEEE Ant. Wirel. Propag. Lett.* 23 (6) (2024) 1739–1743.
- [36] E.H. Houssein, M.E. Hosney, D. Oliva, E.M.G. Younis, A.A. Ali, W.M. Mohamed, An efficient discrete rat swarm optimizer for global optimization and feature selection in chemoinformatics (paper no), *Knowl. -Based Syst.* 275 (2023) 110697.
- [37] S. Kumar, B.S. Yildiz, P. Mehta, N. Panagant, S.M. Sait, S. Mirjalili, A.R. Yildiz, Chaotic marine predators algorithm for global optimization of real-world engineering problems (paper no), *Knowl. -Based Syst.* 261 (2023) 110192.
- [38] Q. Zhang, H. Gao, Z.H. Zhan, J. Li, H. Zhang, Growth Optimizer: A powerful metaheuristic algorithm for solving continuous and discrete global optimization problems (paper no), *Knowl. -Based Syst.* 261 (2023) 110206.
- [39] R.R. Mostafa, M.A. Gaheen, M.A. ElAziz, M.A. Al-Betar, A.A. Ewees, „An improved gorilla troops optimizer for global optimization problems and feature selection, *Knowl. -Based Syst.* 269 (paper 110462) (2023).
- [40] M. Abdel-Salam, A.I. Alzahrani, F. Alblehai, R. Abu Zitar, L. Abualigah, „An improved genghis khan optimizer based on enhanced solution quality strategy for global optimization and feature selection problems, *Knowl. -Based Syst. Pap.* 112237 (2024).

- [41] N.K. Hussein, M. Qaraad, A.M. El Najjar, M.A. Farag, M.A. Elhosseini, S. Mirjalili, D. Guinovart, Schrödinger optimizer: A quantum duality-driven metaheuristic for stochastic optimization and engineering challenges (paper no), *Knowl. -Based Syst.* 328 (2025) 114273.
- [42] E. Aswini, C. Vijayakumaran, Citrus plant classification using gated fusion adaptive graph neural network with harbor seal whiskers optimization algorithm (paper no), *Knowl. -Based Syst.* 329 (2025) 114093.
- [43] R.B. Wang, et al., The animated oat optimization algorithm: a nature-inspired metaheuristic for engineering optimization and a case study on wireless sensor networks (paper no), *Knowl. -Based Syst.* 318 (2025) 113589.
- [44] G. Ativor, V.A. Temeng, Y.Y. Ziggah, Optimisation of multilayer perceptron neural network using five novel metaheuristic algorithms for the prediction of wear of excavator bucket teeth (paper no), *Knowl. -Based Syst.* 321 (2025) 113753.
- [45] S. Selvabharathi, K.S. Dhanalakshmi, P. Prabhu, Adaptive neurospider-har: bio-inspired neural optimization for high-precision human activity recognition (paper no), *Knowl. -Based Syst.* (2025) 115245.
- [46] A.M. Aladdin, T.A. Rashid, LEO: Lagrange elementary optimization, *Neural Comput. Appl.* 37 (2025) 14365–14397.
- [47] A.K. Faraj, A.M. Aladdin, A.A. Ameen, EESB-FDO: enhancing the fitness-dependent optimizer through a modified boundary handling mechanism (Art. no), *Front. Appl. Math. Stat.* 11 (2025) 1640044.
- [48] M. Braik, A. Hammouri, J. Atwan, M.A. Al-Betar, M.A. Awadallah, „White Shark Optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems (paper no), *Knowl. -Based Syst.* 243 (2022) 108457.
- [49] E.H. Houssein, M.E. Hosney, D. Oliva, E.M.G. Younis, A.A. Ali, W.M. Mohamed, An efficient discrete rat swarm optimizer for global optimization and feature selection in chemoinformatics, *Knowl. -Based Syst. Vol.* 275 Pap. no (2023) 110697.
- [50] S. Kumar, B.S. Yildiz, P. Mehta, N. Panagant, S.M. Sait, S. Mirjalili, A.R. Yildiz, Chaotic marine predators algorithm for global optimization of real-world engineering problems, *Knowl. -Based Syst. Vol.* 261 Pap. no (2023) 110192.
- [51] Q. Zhang, H. Gao, Z.H. Zhan, J. Li, H. Zhang, Growth Optimizer: A powerful metaheuristic algorithm for solving continuous and discrete global optimization problems, *Knowl. -Based Syst. Vol.* 261 Pap. no. (2023) 110206.
- [52] R.R. Mostafa, M.A. Gaheen, M.A. ElAziz, M.A. Al-Betar, A.A. Ewees, „An improved gorilla troops optimizer for global optimization problems and feature selection, *Knowl. -Based Syst.* 269 (paper 110462) (2023).
- [53] M. Abdel-Salam, A.I. Alzahrani, F. Alblehai, R. Abu Zitar, L. Abualigah, „An improved genghis khan optimizer based on enhanced solution quality strategy for global optimization and feature selection problems (paper no), *Knowl. -Based Syst.* (2024) 112237.
- [54] M. Verdaasdonk, A. Bart Smolders, A.C.F. Reniers, „Uncertainties introduced by the probe in millimeter-wave planar near-field measurements, *IEEE Trans. Ant. Propag.* 72 (6) (2024) 4862–4873.
- [55] A. Pietrenko-Dabrowska, S. Koziel, M. Al-Hasan, „Expedited yield optimization of narrow- and multi-band antennas using performance-driven surrogates, *IEEE Access.* (2020) 143104–143113.
- [56] J. Wang, X.S. Yang, B.Z. Wang, Efficient gradient-based optimisation of pixel antenna with large-scale connections, *IET Microw. Ant. Prop.* 12 (3) (2018) 385–389.
- [57] J. Dou, H. Gong, H. Chen, Y. Chen, T. Shen, J. Song, A reinforcement learning-based image mapping-topology optimization for antenna design, *IEEE Ant. Wirel. Propag. Lett.* 24 (7) (2025) 1759–1763.
- [58] S. Koziel, A. Pietrenko-Dabrowska, Variable-fidelity simulation models and sparse gradient updates for cost-efficient optimization of compact antenna input characteristics, *Sensors* 19 (8) (2019).
- [59] A. Pietrenko-Dabrowska, S. Koziel, Computationally-efficient design optimization of antennas by accelerated gradient search with sensitivity and design change monitoring, „*IET Microw. Ant. Prop.* 14 (2) (2020) 165–170.
- [60] S. Koziel, A. Pietrenko-Dabrowska, „Reduced-cost electromagnetic-driven optimization of antenna structures by means of trust-region gradient-search with sparse Jacobian updates, „*IET Microw. Ant. Prop.* 13 (10) (2019) 1646–1652.
- [61] F. Feng, J. Zhang, W. Zhang, Z. Zhao, J. Jin, Q. Zhang, Coarse- and fine-mesh space mapping for EM optimization incorporating mesh deformation, *IEEE Microw. Wirel. Comp. Lett.* 29 (8) (2019) 510–512.
- [62] L.-L. Wang, X.-S. Yang, Adjoint sensitivity analysis based on well-conditioned asymptotic waveform evaluation for broadband antenna topology optimization acceleration, *IEEE Trans. Ant. Propag.* 73 (1) (2025) 87–95.
- [63] Y. Wu, et al., An optimized multiband antenna for UWB ad hoc networks based on topology optimization theory, *IEEE Trans. Ant. Propag.* 72 (5) (2024) 3896–3911.
- [64] F. Arndt, WASP-NET: recent advances in fast EM CAD and optimization of waveguide components, feeds and aperture antennas, Chicago, IL, USA, 8–14 July, *IEEE Int. Symp. Ant. Propag.* (2012) 1–2. Chicago, IL, USA, 8–14 July.
- [65] J. Dou, et al., A simulation-driven surrogate parallel improved AGA method for the automated design of antenna, *IEEE Ant. Wirel. Propag. Lett.* 24 (3) (2025) 721–725.
- [66] S. Koziel, A. Pietrenko-Dabrowska, „Expedited feature-based quasi-global optimization of multi-band antennas with Jacobian variability tracking, „*IEEE Access.* 8 (2020) 83907–83915.
- [67] A. Pietrenko-Dabrowska, S. Koziel, Reliable surrogate modeling of antenna input characteristics by means of domain confinement and principal components, *Electronics* 9 (5) (2020) 1–16.
- [68] J. Fu, J. Jin, J. Yang, Z. Liu, J. Qian, H. Lin, Feature-assisted neural network surrogate-based multiphysics optimization for microwave filters, *IEEE Microw. Wirel. Techn. Lett.* 34 (5) (2024) 474–477.
- [69] F. Feng, C. Zhang, W. Na, J. Zhang, W. Zhang, Q. Zhang, Adaptive feature zero assisted surrogate-based EM optimization for microwave filter design, *IEEE Microw. Wirel. Comp. Lett.* 29 (1) (2019) 2–4.
- [70] C. Zhang, F. Feng, V. Gongal-Reddy, Q.J. Zhang, J.W. Bandler, „Cognition-driven formulation of space mapping for equal-ripple optimization of microwave filters, „*IEEE Trans. Microw. Theory Techn.* 63 (7) (2015) 2154–2165.
- [71] W. Chen, Q. Wu, J. Wei, C. Yu, H. Wang, W. Hong, Knowledge-guided and machine-learning-assisted synthesis for series-fed microstrip antenna arrays using base element modeling, *IEEE Trans. Ant. Propag.* 72 (2) (2024) 1497–1509.
- [72] M. Xi, et al., A machine-learning-assisted antenna array isolation optimization method using intelligent surfaces for bistatic in-band full-duplex communication systems, *IEEE Trans. Ant. Propag.* 73 (2) (2025) 962–974.
- [73] J. Li, G. Sun, L. Duan, Q. Wu, Multi-objective optimization for UAV swarm-assisted IoT with virtual antenna arrays, *IEEE Trans. Mob. Comp.* 23 (5) (2024) 4890–4907.
- [74] M. Mahmood, A. Koc, R. Morawski, T. Le-Ngoc, „Achieving capacity gains in practical full-duplex massive MIMO systems: a multi-objective optimization approach using hybrid beamforming, „*IEEE Open. J. Comm. Soc.* 5 (2024) 2268–2286.
- [75] M. Rossi, A. Dierck, H. Rogier, D. Vande Ginste, A stochastic framework for the variability analysis of textile antennas, *IEEE Trans. Ant. Prop.* 62 (16) (2014) 6510–6514.
- [76] S. Koziel and T. Haq, „Computationally-efficient statistical design and yield optimization of resonator-based notch filters by feature-based surrogates, „vol. 13, art. no. 14823, *Sc. Rep.*, 2023.
- [77] S. Koziel, A. Pietrenko-Dabrowska, „Tolerance optimization of antenna structures by means of response feature surrogates, „*IEEE Trans. Ant. Propag.* 70 (11) (2022) 10988–10997.
- [78] J.P. Jacobs, Characterization by Gaussian processes of finite substrate size effects on gain patterns of microstrip antennas, *IET Microw. Ant. Prop.* 10 (11) (2016) 1189–1195.
- [79] J. Cai, J. King, C. Yu, J. Liu, L. Sun, Support vector regression-based behavioral modeling technique for RF power transistors, *IEEE Microw. Wirel. Comp. Lett.* 28 (5) (2018) 428–430.
- [80] Q. Zhou, Y. Wang, P. Jiang, X. Shao, S.-K. Choi, J. Hu, L. Cao, X. Meng, An active learning radial basis function modeling method based on self-organization maps for simulation-based design problems, *Knowl. -Based Syst.* 131 (2017) 10–27.
- [81] J. Dong, W. Qin, M. Wang, Fast multi-objective optimization of multi-parameter antenna structures based on improved BPNN surrogate model, *IEEE Access.* 7 (2019) 77692–77701.
- [82] S. Koziel, N. Calik, P. Mahouti, M.A. Belen, Accurate modeling of antenna structures by means of domain confinement and pyramidal deep neural networks, *IEEE Trans. Ant. Prop.* 70 (3) (2022) 2174–2188.
- [83] D. Gorissen, K. Crombecq, I. Couckuyt, T. Dhaene, P. Demeester, A surrogate modeling and adaptive sampling toolbox for computer based design, *J. Mach. Learn. Res.* 11 (2010) 2051–2055.
- [84] S. Marelli, B. Sudret, UQLab: a framework for uncertainty quantification in Matlab, 2nd Int. Conf. Vulnerability Risk Anal. Manag. (ICVRAM 2014) Univ. Lond. UK July 13–15 (2014) 2554–2563.
- [85] M.A. Bouhellel, J.T. Hwang, N. Bartoli, R. Lafage, J. Morlier, J.R.R.A. Martins, A Python surrogate modeling framework with derivatives, *Adv. Eng. Soft.* (2019) 102662.
- [86] Ver Matlab, the MathWorks Inc, Natick, Massachusetts, USA, 2025, p. 2025.
- [87] N. Taran, D.M. Ionel, D.G. Dorrell, Two-level surrogate-assisted differential evolution multi-objective optimization of electric machines using 3-D FEA, paper 8107605, *IEEE Trans. Magn.* 54 (11) (Nov. 2018). paper 8107605.
- [88] Q. Wu, W. Chen, C. Yu, H. Wang, W. Hong, Multilayer machine learning-assisted optimization-based robust design and its applications to antennas and arrays, *IEEE Trans. Ant. Prop.* 69 (9) (2021) 6052–6057.
- [89] S. Koziel, A. Pietrenko-Dabrowska, S. Szczepanski, „Multi-objective artificial-intelligence-based parameter tuning of antennas using variable-fidelity machine learning, *Sc. Rep.* 15 (paper 21693) (2025).
- [90] M.S. Yahya, S. Soeung, S.K.A. Rahim, U. Musa, S.S. Ba Hashwan, M.A. Haque, Machine learning-optimized compact frequency reconfigurable antenna with RSSI enhancement for long-range applications, *IEEE Access.* 12 (2024) 10970–10987.
- [91] H. Huang, X.-S. Yang, Y.-Y. Yang, B.-Z. Wang, An effective antenna shape optimization framework employing machine-learning-based approach, *IEEE Trans. Ant. Propag.* 73 (10) (2025) 8360–8365.
- [92] Y.-F. Liu, L.-Y. Xiao, Q. Huo Liu, Machine learning-based design scheme for multifunctional antenna arrays with reconfigurable scattering patterns, *IEEE Trans. Ant. Propag.* 73 (7) (2025) 4535–4548.
- [93] P. Gu, Z. Cao, Z. He, D. Ding, Design of ultrawideband RCS reduction metasurface using space mapping and phase cancellation, *IEEE Ant. Wirel. Prop. Lett.* 22 (6) (2023) 1386–1390.
- [94] Z.-Y. Zhang, et al., Circuit model for nth-order stacked patch antenna, parameter extraction, and radiation prediction, *IEEE Trans. Ant. Prop.* 72 (2) (2024) 1329–1342.
- [95] J. Zhang, X. Yan, H. Luo, Y. Guo, An efficient transistor-model-assisted layout synthesis approach using improved implicit space mapping for high-performance MMIC PAs, *IEEE Trans. Circuits Syst. II Express Briefs* 71 (5) (2024) 2639–2643.
- [96] S. Koziel, S.D. Unnsteinsson, Expedited design closure of antennas by means of trust-region-based adaptive response scaling, *IEEE Ant. Wirel. Propag. Lett.* 17 (6) (June 2018) 1099–1103.

- [97] Y. He, J. Huang, W. Li, L. Zhang, S.W. Wong, Z.N. Chen, Hybrid method of artificial neural network and simulated annealing algorithm for optimizing wideband patch antennas, *IEEE Trans. Antennas Propag.* 72 (1) (2024) 944–949.
- [98] A. Touhami, S. Collardey, A. Shariha, A global optimization method for wideband and small supergain arrays design using Artificial Neural Network, *IEEE Open. J. Ant. Propag.* 4 (2023) 1016–1028.
- [99] Y. Liu, et al., An efficient method for antenna design based on a self-adaptive Bayesian neural network-assisted global optimization technique optimization, *IEEE Trans. Antennas Propag.* 70 (12) (2022) 11375–11388.
- [100] A. Sonker, A.K. Nayak, T. Goel, A. Patnaik, Multifunctional antenna design for wireless consumer electronic devices: a soft-computing approach, *IEEE Can. J. Electr. Comput. Eng.* 46 (2) (2023) 144–156.
- [101] Y. Liu, P. Chen, J. Tian, J. Xiao, S. Noghianian, Q. Ye, Hybrid ANN-GA optimization method for minimizing the coupling in MIMO antennas (article no.), *AEU Int. J. Electron. Comm.* 175 (2024) 155068 (article no.).
- [102] Z.Z. Stanković, D.I. Olčan, N.S. Dončov, B.M. Kolundžija, “Consensus deep neural networks for antenna design and optimization,” *IEEE Trans. Antennas Propag.* 70 (7) (2022) 5015–5023.
- [103] K. Yasmeen, K.V. Mishra, A.V. Subramanyam, S.S. Ram, Circularly polarized Fabry–Pérot cavity sensing antenna design using generative model, pp. 1–4, *Art no*, *IEEE Sens. Lett.* 7 (2) (2023) 3500304.
- [104] A. Gupta, E.A. Karahan, C. Bhat, K. Sengupta, U.K. Khankhoje, Tandem neural network based design of multiband antennas, *IEEE Trans. Antennas Propag.* 71 (8) (2023) 6308–6317.
- [105] L. Kouhalvandi, L. Matekovits, Hyperparameter optimization of long short-term memory-based forecasting DNN for antenna modeling through stochastic methods, *IEEE Antennas Wirel. Propag. Lett.* 21 (4) (2022) 725–729.
- [106] J. Tan, Y. Shao, J. Zhang, J. Zhang, Efficient antenna modeling and optimization using multifidelity stacked neural network, *IEEE Trans. Antennas Propag.* 72 (5) (2024) 4658–4663.
- [107] J. Li, et al., An adaptive evolutionary neural network-based optimization design method for wideband dual-polarized antennas, *IEEE Trans. Antennas Propag.* 71 (10) (2023) 8165–8172.
- [108] G. Qi, S. Bao, Y. Wei, Y. Zhang, Accurate antenna design by deep auto-encoder surrogate model assisted particle swarm optimization, *IEEE 5th Int. Conf. Electron. Inf. Comm. Technol. (ICEICT) Hefei China 2022* (2022) 875–880.
- [109] J.P. Liu, B.Z. Wang, C.S. Chen, R. Wang, Inverse design method for horn antennas based on knowledge-embedded physics-informed neural networks, *IEEE Antennas Wirel. Propag. Lett.* 23 (6) (2024) 1665–1669.
- [110] Y.F. Liu, L. Peng, W. Shao, An efficient knowledge-based artificial neural network for the design of circularly polarized 3-D-printed lens antenna, *IEEE Trans. Antennas Propag.* 70 (7) (2022) 5007–5014.
- [111] L. Guo, W. Cao, W. Hu, Z. Lu, M. Wu, Attention-modulated feature fusion neural network for inverse modeling of microwave filters, *IEEE Microw. Wirel. Techn. Lett.* 35 (4) (2025) 376–379.
- [112] M. Sedaghat, R. Trinchero, Z.H. Firouzeh, F.G. Canavero, Compressed machine learning-based inverse model for design optimization of microwave components, *IEEE Trans. Microw. Theory Techn.* 70 (7) (2022) 3415–3427.
- [113] L. Yuan, L. Wang, X.-S. Yang, H. Huang, B.-Z. Wang, An efficient artificial neural network model for inverse design of metasurfaces, *IEEE Ant. Wirel. Propag. Lett.* 20 (6) (2021) 1013–1017.
- [114] J. Jin, F. Feng, Q.-J. Zhang, An overview of neural network techniques for microwave inverse modeling, *IEEE Int. Symp. Radio. -Freq. Integr. Techn. (RFIT) Hualien Taiwan*. (2021) 1–2.
- [115] M.R. Khan, C.L. Zekios, S. Bhardwaj, S.V. Georgakopoulos, A deep learning convolutional neural network for antenna near-field prediction and surrogate modeling, *IEEE Access.* 12 (2024) 39737–39747.
- [116] Y. Su, et al., Time-domain scattering parameters-based neural network inverse model for antenna designs, *IEEE Antennas Wirel. Propag. Lett.* 23 (7) (2024) 1976–1980.
- [117] J. Jin, Q. Su, Y. Xu, Z. He, Y. Lu, Efficient radiation pattern prediction of array antennas based on complex-valued graph neural networks, *IEEE Antennas Wirel. Propag. Lett.* 21 (12) (2022) 2467–2471.
- [118] Z. Lv, L. Wang, Z. Han, J. Zhao, W. Wang, Surrogate-assisted particle swarm optimization algorithm with Pareto active learning for expensive multi-objective optimization, *IEEE/CAA J. Autom. Sin.* 6 (3) (May. 2019) 838–849.
- [119] R. Hu, V. Monebhurrin, R. Himeno, H. Yokota, F. Costen, An adaptive least angle regression method for uncertainty quantification in FDTD computation, *IEEE Trans. Ant. Prop.* 66 (12) (2018) 7188–7197.
- [120] A.C. Yücel, H. Bağcı, E. Michielssen, An ME-PC enhanced HDMR method for efficient statistical analysis of multiconductor transmission line networks, *IEEE Trans. Comp. Packag. Manuf. Techn.* Vol. 5 (5) (2015) 685–696.
- [121] M.T.H. Kaib, A. Kouadri, M.-F. Harkat, A. Bensmail, M. Mansouri, Improvement of kernel principal component analysis-based approach for nonlinear process monitoring by data set size reduction using class interval, *IEEE Access.* 12 (2024) 11470–11480.
- [122] J.P. Jacobs, S. Koziel, Two-stage framework for efficient Gaussian process modeling of antenna input characteristics, *IEEE Trans. Antennas Prop.* 62 (2) (2014) 706–713.
- [123] J. Li, A. Yang, C. Tian, L. Ye, B. Chen, Multi-fidelity Bayesian algorithm for antenna optimization, *J. Syst. Eng. Electr.* 33 (6) (2022) 1119–1126.
- [124] S. Koziel, A. Pietrenko-Dabrowska, “Performance-based nested surrogate modeling of antenna input characteristics,” *IEEE Trans. Ant. Prop.* 67 (5) (2019) 2904–2912.
- [125] S. Koziel, A. Pietrenko-Dabrowska, *Performance-driven surrogate modeling of high-frequency structures*, Springer, New York, 2020.
- [126] S. Koziel, A. Pietrenko-Dabrowska, “Knowledge-based performance-driven modeling of antenna structures,” *Knowl. Based Syst. Pap.* 107698 (2021).
- [127] A. Pietrenko-Dabrowska, S. Koziel, “Antenna modeling using variable-fidelity EM simulations and constrained co-kriging,” *IEEE Access.* 8 (1) (2020) 91048–91056.
- [128] A. Pietrenko-Dabrowska, S. Koziel, M. Al-Hasan, “Cost-efficient bi-layer modeling of antenna input characteristics using gradient kriging surrogates,” *IEEE Access.* 8 (2020) 140831–140839.
- [129] S. Koziel, P. Mahouti, N. Calik, M.A. Belen, S. Szczepanski, Improved modeling of miniaturized microwave structures using performance-driven fully-connected regression surrogate, *IEEE Access.* 9 (2021) 71470–71481.
- [130] T. Iye, P. van Wyk, T. Matsumoto, Y. Susukida, S. Takaya, Y. Fujii, Neural network-based phase estimation for antenna array using radiation power pattern, *IEEE Ant. Wirel. Propag. Lett.* 21 (7) (2022) 1348–1352.
- [131] R.G. Ayestaran, F. Las-Heras, L.F. Herran, High-accuracy neural-network-based array synthesis including element coupling, *IEEE Ant. Wirel. Propag. Lett.* 5 (2006) 45–48.
- [132] P. Liu, D. Liu, Y. Li, S. Ye, D. Su, Attention-based ResNet for radiation pattern prediction of phased array antenna, *IEEE Ant. Wirel. Propag. Lett.* 23 (12) (2024) 4453–4457.
- [133] G.D. King, M.Asaduzzaman Towfiq, A.C. Gurbuz, B.A. Cetiner, Beam coefficient prediction for antenna arrays using physics-aware convolutional neural networks, *IEEE Access.* 12 (2024) 176908–176919.
- [134] Z. Zhou, Z. Wei, J. Ren, Y. Yin, G.F. Pedersen, M. Shen, Two-order deep learning for generalized synthesis of radiation patterns for antenna arrays, *IEEE Trans. Ant. Intel.* 4 (5) (2023) 1359–1368.
- [135] P. Singh, S.S. Panda, J.C. Dash, B. Riscob, S.K. Pathak, R.S. Hegde, Rapid multi-objective antenna synthesis via deep neural network surrogate-driven evolutionary optimization, *IEEE J. Multiscale Multiphysics Comp. Techn.* 10 (2025) 151–159.
- [136] T.N. Kapetanakis, I.O. Vardiambasis, E.I. Lourakis, A. Maras, Applying neuro-fuzzy soft computing techniques to the circular loop antenna radiation problem, *IEEE Ant. Wirel. Propag. Lett.* 17 (9) (2018) 1673–1676.
- [137] D. Pujara, A. Modi, N. Pisharody, J. Mehta, Predicting the performance of pyramidal and corrugated horn antennas using ANFIS, *IEEE Ant. Wirel. Propag. Lett.* 13 (2014) 293–296.
- [138] M.R. Khan, C.L. Zekios, S. Bhardwaj, S.V. Georgakopoulos, A deep learning convolutional neural network for antenna near-field prediction and surrogate modeling, *IEEE Access.* 12 (2024) 39737–39747.
- [139] S. Jan, Y. Liu, C.D. Sarris, Efficient neural network-based reconstruction of three-dimensional antenna radiation patterns from two-dimensional cuts, *IEEE Open. J. Ant. Propag.* 6 (3) (2025) 726–734.
- [140] L.-Y. Xiao, W. Shao, F.-L. Jin, B.-Z. Wang, Q.H. Liu, Inverse artificial neural network for multiobjective antenna design, *IEEE Trans. Ant. Propag.* 69 (10) (2021) 6651–6659.
- [141] L.-Y. Xiao, W. Shao, F.-L. Jin, B.-Z. Wang, Multiparameter modeling with ANN for antenna design, *IEEE Trans. Ant. Propag.* 66 (7) (2018) 3718–3723.
- [142] D. Gorissen, K. Crombecq, I. Couckuyt, T. Dhaene, P. Demeester, A surrogate modeling and adaptive sampling toolbox for computer based design, *J. Mach. Learn. Res.* 11 (2010) 2051–2055.
- [143] X.R. Li, Z. Zhao, Evaluation of estimation algorithms part I: comprehensive measures of performance, *IEEE Trans. Aerosp. Electr. Syst.* 42 (4) (2006) 1340–1358.
- [144] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, [Online]. Available (arXiv.), *Going Deep. convolutions* (Sep-2014) 17, <https://doi.org/10.48550/arXiv.1409.4842>.
- [145] A. Vaswani, N. Shazeer, N. Parmar, L. Uszkoreit, J. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, arXiv preprint, arXiv:1706.03762, *Atten. Is. all you need* (Jun. 2017).
- [146] N. Shazeer, A. Mirhoseini, K. Maziarz, A. Davis, Q. Le, G. Hinton, J. Dean, “Outrageously large neural networks: The sparsely-gated mixture-of-experts layer,” *arXiv preprint arXiv:1701.06538* (Available), [Online] (Jan. 2017), <https://doi.org/10.48550/arXiv.1701.06538>.
- [147] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, *Int. Conf. Learn. Represent. (ICLR)* (2015). <https://arxiv.org/abs/1412.6980>.
- [148] B. Beachkofski and R. Grandhi, “Improved distributed hypercube sampling,” *American Institute of Aeronautics and Astronautics*, paper AIAA 2002-1274, 2002.
- [149] M.G.N. Alsath, M. Kanagasabai, Compact UWB monopole antenna for automotive communications, *IEEE Trans. Ant. Prop.* 63 (9) (2015) 4204–4208.
- [150] S.D. Unnsteinsson, S. Koziel, Generalized Pareto ranking bisection for computationally feasible multi-objective antenna optimization, *Int. J. RF & Microw. CAE* 28 (8) (2018).
- [151] A.L.J. Forrester, A.J. Keane, Recent advances in surrogate-based optimization, *Prog. Aerosp. Sci.* 45 (2009) 50–79.
- [152] C.K.I. Williams, C.E. Rasmussen, Gaussian processes for regression, in: D. Touretzky, M.C. Mozer, M. Hasselmo (Eds.), *Advances in Neural Information Proc. Syst.*, 8, MIT Press, 1995.
- [153] M. Awad, R. Khanna, *Support vector regression. Efficient Learning Machines*, Apress, Berkeley, CA, 2015.
- [154] R. Vang-Mata (Ed.), *Multilayer perceptrons*, Nova Science Pub. Inc, 2020.
- [155] F.M. Salem, *Recurrent Neural Networks: From Simple to Gated Architectures*, Springer, New York, 2023.
- [156] C.C. Agrawal, *Neural Networks and Deep Learning: A Textbook*, 2nd Ed, Springer, New York, 2023.

- [157] Z. Lv, L. Wang, Z. Han, J. Zhao, W. Wang, Surrogate-assisted particle swarm optimization algorithm with Pareto active learning for expensive multi-objective optimization, *IEEE/CAA J. Autom. Sin.* 6 (3) (2019) 838–849.
- [158] A.R. Conn, N.I.M. Gould, P.L. Toint, *MPS-SIAM Series on Optimization*, Trust-Reg. Methods (2000).
- [159] R.K. Muhammed, A.M. Aladdin, R.K. Hamad, D.R. Hamad, T.A. Rashid, Fundamentals of Quantum-Inspired Neural Computing. *Quantum-Inspired Neural Networks: Future Perspectives and Challenges*, CRC Press, 2026, pp. 1–29, <https://doi.org/10.1201/9781003682950-1>.