(Received 19 June 2025, Accepted 27 August 2025, Scheduled 8 September 2025)

Multi-Mechanism Fusion Based 1D U-Net Models for Antenna Forward and Inverse Design

Ximin Yang[1], Jingchang Nan[1,*], and Minghuan Wang[2]

[1]School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China
[2]School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210000, China

ABSTRACT: This study investigates the relation between the physical parameters and scattering parameter (S_{11}) curves of antennas, and proposes two deep-neural-network-based frameworks respectively for antenna forward and inverse designs, improving the design efficiency compared to the conventional electromagnetic (EM) simulation approaches. In this study, a one-dimensional (1D) U-Net is utilized as the backbone of the two models and is enhanced with multiple mechanisms — diffusion mechanism, channel attention, and spatial attention. Therefore, the models more effectively capture the sequential features of data. In the forward design, the model quickly predicts the S_{11} curves from given physical parameters with an accuracy improvement of at least 63% RMSE and 70% MAE compared to the improved one-dimensional convolutional neural network (1D-MCNN) and deep multi-layer perceptron (DMLP), thus realizing the surrogate model of conventional methods to some extent. In the inverse design, another model directly infers the physical parameters corresponding to the target S_{11} curves with an accuracy improvement of at least 21% RMSE and 38% MAE compared to the baseline models (1D U-Net and MLP), thereby eliminating the iterative process of traditional methods and accelerating the antenna design. The experimental results demonstrate the significant advantages of the proposed deep neural network frameworks in terms of accuracy and efficiency for both forward and inverse designs of antennas, offering a powerful alternative to conventional electromagnetic simulation-based approaches.

1. INTRODUCTION

The conventional design of antenna typically relies on computer-aided design (CAD) software that accurately simulates the electromagnetic (EM) field [1]. However, these traditional methods often require high level of expertise, including extensive EM knowledge and sufficient design experience. Additionally, the necessity for numerous EM simulations leads to significant computational and time costs in practical design. Consequently, all these disadvantages have impeded the expeditious advancement of antenna technology [2].

Deep learning (DL) has been successfully applied in various fields [3–5] and has emerged as a powerful tool for antenna design in recent years [6–8] because of its exceptional ability to learn complex and nonlinear relations among various parameters. As a cornerstone of DL, a variety of deep neural networks (DNNs) have been effectively utilized as alternative solutions to traditional methods in antenna research.

DNNs construct mapping models between physical parameters and EM responses through data training. The application of DNNs in antenna design is primarily categorized into forward design and inverse design [9]. In forward models, physical parameters serve as inputs, enabling accurate and rapid predictions of EM responses. These predictions can, to some extent, act as surrogates for traditional antenna simulations [10–16].

The S_{11} parameter is one of the most critical scattering parameters (S-parameters) for evaluating an antenna's input-port matching. Its magnitude (usually expressed in dB) directly corresponds to the power reflection coefficient between the antenna and the feed line. This parameter is essential for determining the antenna's operational bandwidth, such as the frequency range over which $|S_{11}| \leq -10$ dB. Therefore, many studies focus on predicting S_{11} curves utilizing DNNs such as deep multi-layer perceptron (DMLP) [10], one-dimensional convolutional neural network (1D-MCNN) [11], Image-CNN-LSTM [2], and convolutional neural networks (CNNs) based on machine vision technologies [12, 13], among others. Ninkovic et al. [14] proposed a mixture of experts (MoE) method for antenna modeling and compared its accuracy and training time to single DNNs and ensembles of DNNs.

Moreover, forward design has been employed in the design and optimization of antenna arrays. Liu et al. [15] developed a residual network (ResNet) incorporating an attention mechanism to predict the radiation patterns of phased array (PPAs). Jin et al. [16] proposed a complex-valued graph neural network (GNN) with residual connections to efficiently predict the radiation patterns for antenna arrays with different geometric structures.

Forward design has been extensively and successfully employed in antenna design and optimization. Nevertheless, it remains an iterative process that involves a certain degree of trial and error, making it impractical to completely eliminate tedious optimization efforts [9].

* Corresponding author: Jingchang Nan (nanjingchang@lntu.edu.cn).

In contrast, the inverse design employs EM responses as the input and physical parameters as the output. A trained inverse model can directly infer the physical parameters corresponding to the target EM responses, making the design process more streamlined and efficient [17, 18].

Taking S-parameters or radiation patterns as the input, the required antenna structures can be obtained at once through inverse models, including the multilayer perceptron neural network (MLP-NN) inverse model based on time-domain scattering (TDS) parameters (TDS-MLP-NN) [9], and deep CNN for modeling template-free EM structures, among others. Liu et al. [19] proposed a knowledge-embedded PINNs (KE-PINNs) inverse design framework for horn antenna design, which effectively achieves super-gain and beam deflection, respectively. Shereen et al. [20] proposed a data-driven DL model capable of learning the implicit mapping between antenna geometries and their electromagnetic responses, achieving bidirectional prediction of either geometry or performance in real-time.

Yao et al. [21] put forward a fault detection method for array antenna units based on a deep CNN (DConvNet), which took the radiation pattern as input, thereby achieving fault diagnosis of antenna arrays in a complex environment. Chen et al. [22] introduced a deep learning-based method using ControlNet diffusion model to design reflectarray antennas, achieving a maximum gain of 28.0 dBi at 10 GHz with an aperture efficiency of 50.3% and low sidelobe levels in both the E-plane and H-plane.

Furthermore, integrating forward models with intelligent optimization algorithms is also regarded as an inverse design method. For instance, Nan et al. [23] proposed an inverse iteration method for antenna modeling based on back propagation neural network (BPNN) with Levenberg-Marquardt (LM) algorithm and Ant Lion Optimizer (ALO). Karahan et al. [17] came up with an inverse design method based on CNN and evolutionary algorithms, which can be utilized to design nearly arbitrary planar antenna structures.

This study focuses on mapping the relation between the physical parameters of antennas and highly nonlinear S_{11} curves. Taking a UWB stepped microstrip monopole antenna [24] as a case study, we first propose an antenna forward model — 1D ConDiffusion-U-Net, which is based on a one-dimensional (1D) U-Net that integrates conditional and diffusion mechanisms. The forward model serves as a surrogate that can partially replace traditional EM simulation methods to predict the S_{11} curves efficiently.

Then, an inverse model — 1D U-Net-CSA, based on 1D U-Net combined with channel and spatial attention mechanisms — is proposed for antenna inverse design. This model can directly output the physical parameters corresponding to the target S_{11} curves, effectively eliminating the iterative process of traditional methods.

The remainder of this paper is organized as follows. In Section 2, the theoretical foundations and details of the forward model (1D ConDiffusion-U-Net) and inverse model (1D U-Net-CSA) are introduced. Section 3 presents and analyzes the application process and results of the antenna case to indicate the advantages of the proposed models. In Section 4, the conclusions of our comprehensive studies and future works are provided.

2. THEORETICAL FOUNDATIONS AND DETAILS OF PROPOSED MODELS

2.1. Convolutional Neural Network

Convolutional Neural Network (CNN), U-Net [25], and 1D U-Net all employ convolutional operations to extract local features through receptive fields and weight sharing, followed by nonlinear activation functions for complex mapping. These DNNs (CNN, U-Net, 1D U-Net) incorporate both down-sampling and up-sampling pathways to capture trends and detailed information across multiple scales. Nevertheless, traditional CNNs typically terminate with full connection layers or a final CNN layer, propagating information strictly in an unidirectional and deep hierarchy, which can lead to the loss of fine-grained details from shallower layers.

In contrast, U-Net [25] effectively addresses the aforementioned limitations through its symmetric encoder-decoder structure and cross-layer skip-connections. In the encoder, feature maps progressively aggregate global context via multiple down-sampling operations. In the decoder, skip-connections concatenate high-resolution feature maps from the corresponding encoding layers with up-sampled representations, thereby restoring spatial details and facilitating the deep fusion of local and global information. Consequently, U-Net has demonstrated remarkable performance in tasks such as image processing [26, 27] and semantic segmentation [28, 29].

In this study, both proposed models are based on 1D U-Net structures, derived by substituting two-dimensional (2D) convolutions in the original U-Net with 1D convolutions tailored for sequence data. The 1D U-Net retains the strengths of U-Net, including multi-scale feature fusion and skip-connections, which are utilized in 1D signal processing, such as regression-based prediction of antenna EM responses. In the encoder path, the network captures the global trend of full-band S_{11} curves; in the decoder path, skip-connections recover local details around each resonance to ensure sensitivity to small frequency variations. Moreover, reducing convolutional complexity and the total number of parameters mitigates the risk of overfitting while improving computational efficiency and model robustness.

2.2. Details of Proposed Forward Model

The proposed forward model (1D ConDiffusion-U-Net) aims at predicting S_{11} curves of input physical parameters. The architecture of the model is shown in Fig. 1(c). The whole architecture comprises three encoder-decoder levels and a bottleneck convolution.

The forward surrogate is formalized as (1).

$$\hat{S}_{11} = f_\theta(x_T, P) \tag{1}$$

Define $x_0 \in \mathbb{R}^L$ as a clean S_{11} curve sampled at L frequency bins, and $P \in \mathbb{R}^{d_{\text{cond}}}$ denotes the antenna physical parameter vector. The diffusion surrogate models the conditional distribution $p(x_0 \mid P)$ by learning noise residuals ϵ_θ that enable a stepwise denoising procedure. Sampling is initialized from

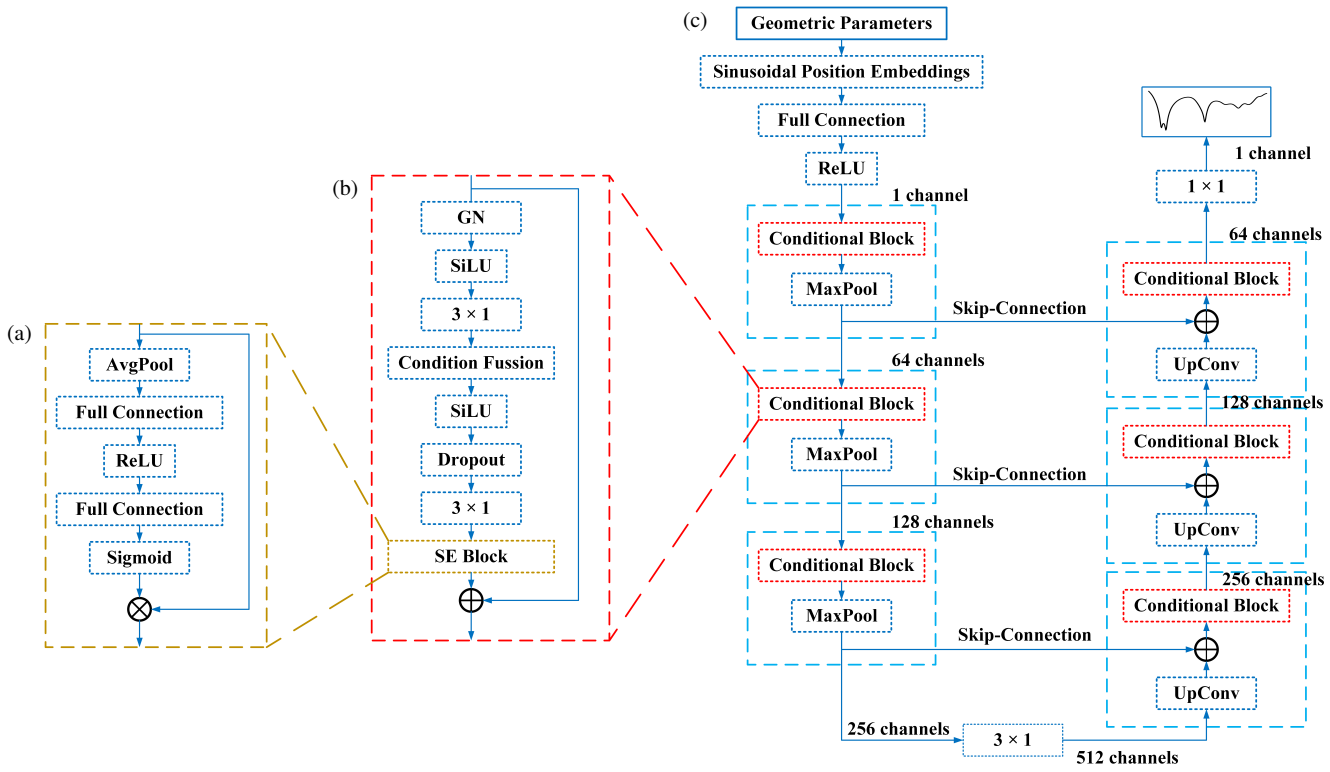

FIGURE 1. Architecture of the forward model based on 1D ConDiffusion-U-Net: (a) structure of the SE block; (b) structure of the conditional block; (c) architecture of 1D ConDiffusion-U-Net.

$x_T \sim \mathcal{N}(0, I)$ and proceeds by iterating reverse updates to obtain a predicted S_{11} conditioned on P.

The forward process is defined by the Markov kernel as (2),

$$\begin{aligned} q(x_t \mid x_{t-1}) &= \mathcal{N}\left(x_t; \sqrt{\alpha_t} x_{t-1}, \beta_t I\right) \\ \alpha_t &= 1 - \beta_t \\ \beta_t &= \beta_{\min} + \frac{t}{T}(\beta_{\max} - \beta_{\min}), \quad t = 1, \ldots, T \\ \beta_{\min} &= 10^{-4}, \quad \beta_{\max} = 2 \times 10^{-2}, \end{aligned} \tag{2}$$

where $t = 1, \ldots, T$ indexes diffusion timesteps, and the cumulative product $\overline{\alpha}_t = \prod_{s=1}^{t} \alpha_s$.

During reverse sampling, the predicted noise residual $\epsilon_\theta(x_t, t, P)$ is used in the closed-form posterior approximation:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \overline{\alpha}_t}} \epsilon_\theta(x_t, t, P) \right) + \sqrt{\beta_t} z \tag{3}$$

where $z \sim \mathcal{N}(0, I)$. Iterating Eq. (3) from $t = T$ down to $t = 1$ yields a sample x_0 interpreted as the S_{11} prediction conditioned on P.

To enable the model to perceive "time-steps" in the diffusion process t, Sinusoidal Positional Embeddings are employed to encode t into a D-dimensional vector via (4), which enables the network to discriminate different diffusion scales across frequencies.

$$\begin{aligned} \mathrm{PE}_{2k}(t) &= \sin(t \cdot \omega_k) \\ \mathrm{PE}_{2k+1}(t) &= \cos(t \cdot \omega_k) \end{aligned}$$

$$\omega_k = \exp\left(\frac{-\ln(10000) k}{(D/2) - 1} \right) \tag{4}$$

Here, k is the index, ranging from 0 to $(D/2) - 1$, with sine and cosine functions applied to the even and odd dimensions, respectively. In this way, the positional encoding enables the model to distinguish inputs at different diffusion scales within the feature space.

Each encoder/decoder stage applies a residual conditional block (Fig. 1(b)) that first performs pre-activation [30] (GroupNorm followed by SiLU) and a 1D convolution to produce an intermediate feature h. Time and condition embeddings are projected by linear layers to produce channel-wise scale and shift vectors; these are combined and applied with LayerNorm via a channel-wise multiply-add operation, i.e., $\gamma = 1 + \gamma_t + \gamma_p$, $\beta = \beta_t + \beta_p$ and $h' = \gamma \odot \mathrm{LayerNorm}(h) + \beta$ [31]. The modulated feature h' is then reweighted by a Squeeze-and-Excitation (SE) [32] (Fig. 1(a)) channel-attention branch computed from global average pooled statistics, and the resulting feature is added to a residual projection of the input (identity or 1×1 convolution when channel dimensions differ). This conditional block therefore integrates pre-activation, FiLM-style conditioning on timestep and antenna parameters, SE channel attention, and a residual connection.

Down-sampling is implemented via MaxPool1d with stride 2 after each encoder stage; the bottleneck at the coarsest scale is a convolutional block of identical internal form to the encoder. The decoder mirrors the encoder: each up-sampling uses ConvTranspose1d to double temporal resolution, concatenates the up-sampled feature with the corresponding encoder feature via

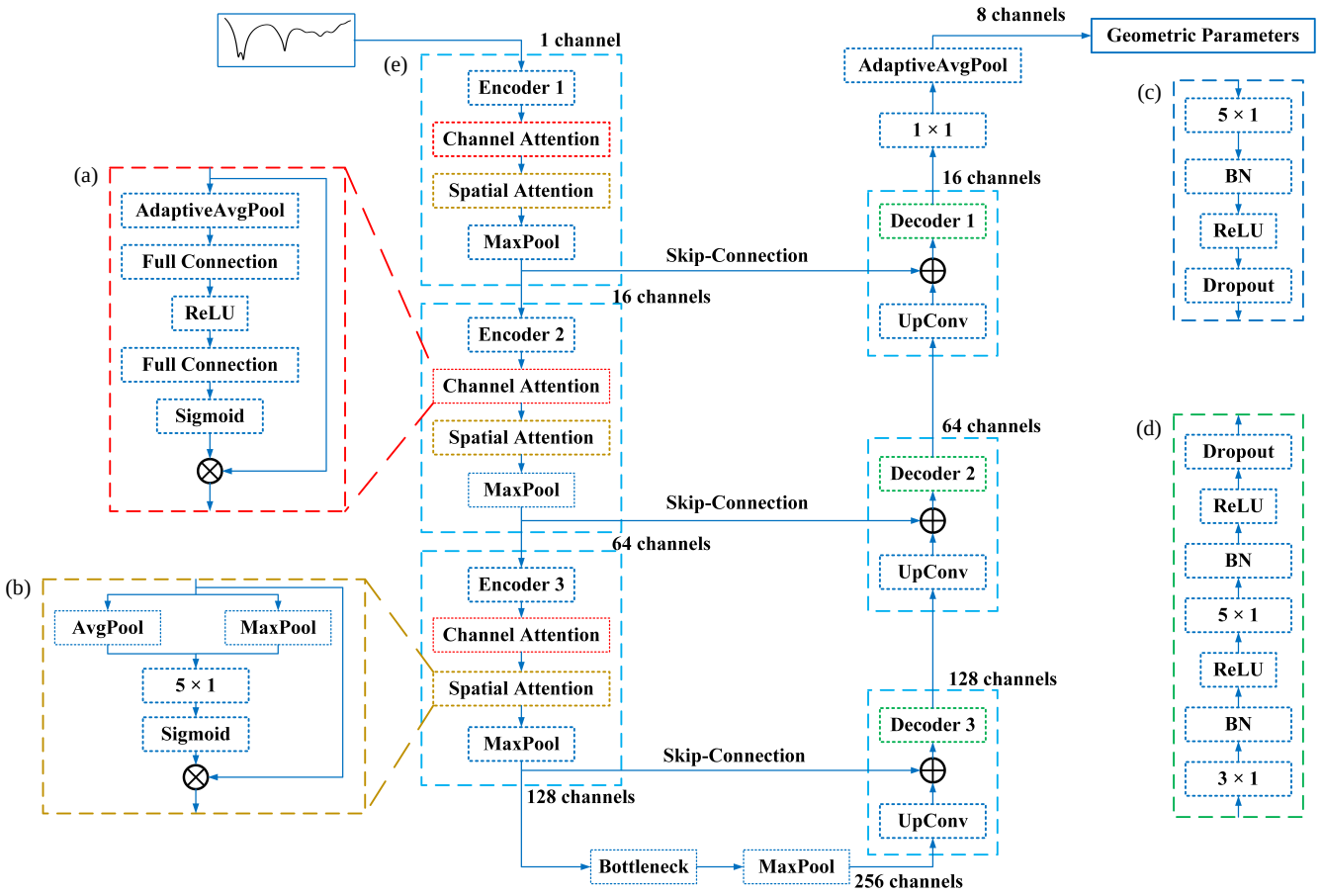

FIGURE 2. Architecture of the inverse model based on 1D U-Net-CSA: (a) structure of the channel attention; (b) structure of the spatial attention; (c) structure of the encoder; (d) structure of the decoder; (e) architecture of 1D U-Net-CSA.

a skip-connection, and refines the concatenated representation with the same conditional block. The network output is produced by a final 1×1 convolution that maps the decoded feature map to the required output channels representing the S_{11} curve.

Notation: L denotes the number of frequency bins; T is the total diffusion timesteps; D is the time-embedding dimension; d_{cond} denotes the dimension of the antenna parameter vector; α_t, β_t, $\overline{\alpha}_t$ are noise-schedule parameters; GAP denotes global average pooling; $\text{Proj}_{1\times1}$ denotes a 1×1 projection convolution; GroupNorm_G, LayerNorm, SiLU, and σ denote standard operations; W_t, W_p, W_1, and W_2 denote trainable linear maps used in FiLM and SE modules.

2.3. Details of Proposed Inverse Model

The proposed inverse model (1D U-Net-CSA) is designed to directly obtain physical parameters of the target S_{11} curves. The architecture of the model is shown in Fig. 2(e). The whole architecture also comprises three encoder-decoder levels and a bottleneck convolution.

The inverse network is defined as (5):

$$\hat{P} = g_\phi(X_0), \tag{5}$$

where $X_0 \in \mathbb{R}^L$ is a target S_{11} curve sampled at L frequency bins, and $\hat{P} \in \mathbb{R}^{d_{\text{cond}}}$ denotes the predicted antenna geometric parameters.

In the down-sampling stage, each level comprises three steps of the Encoder Block (Fig. 2(c)), Channel Attention (Fig. 2(a)), and Spatial Attention (Fig. 2(b)):

- Encoder Block: As (6), the input feature map passes through convolution, BatchNorm, ReLU, and Dropout, producing multi-channel representations.

$$E(x) = \text{Dropout}_{p=0.2}\left[\text{ReLU}(\text{BN}(\text{Conv}_{5,2}(x)))\right], \tag{6}$$

after which channel attention and spatial attention are applied, and the result is down-sampled by MaxPool1d (kernel size 2, stride 2).

- Channel Attention: It implements the SE operation of the forward model above. Global average pooling across time, two full connection layers with ReLU followed by Sigmoid to produce per-channel scaling factors and re-weighting of the feature map.

- Spatial Attention [33]: This module, as (7), concatenates channel-wise average and max pooling along the channel axis to form a two-channel descriptor, applies a convolution and Sigmoid to generate an attention map over the temporal dimen-

sion, and multiplies it back to the input feature map.

$$\begin{aligned} a_{\text{avg}} &= \frac{1}{C}\sum_{i=1}^{C} x_{:,i,:} \\ a_{\max} &= \max_{i=1\ldots C} x_{:,i,:} \\ m &= [a_{\text{avg}};\, a_{\max}] \in \mathbb{R}^{B\times 2\times L} \\ A_s(x) &= x \odot \sigma\left(\text{Conv}_{5,2}(m)\right) \end{aligned} \tag{7}$$

with $\text{Conv}_{5,2}$ a kernel-5, pad-2 conv along L.

Then, more multi-scale information can be obtained through MaxPool1d.

At the coarsest scale, a single convolutional block as the bottleneck which is the same as the Encoder Block (Fig. 2(d)) integrates global context.

After three encoder levels and bottleneck processing (the bottleneck uses the same internal block structure as the encoder), the decoder path performs up-sampling. Each up-sampling stage begins with a transposed 1D convolution that approximately doubles the temporal length; the up-sampled tensor is aligned (center-cropped if necessary) and concatenated with the corresponding encoder feature map via a skip-connection. The concatenated feature is then refined by a decoder block composed of two successive convolutional refinements (first a 1D conv with kernel size 3, padding 1, then a 1D conv with kernel size 5, padding 2), each followed by BatchNorm and ReLU, and finally dropout. In compact form as (8):

$$\begin{aligned} u &= \text{ConvTranspose1d}_{\text{up}}(v) \\ y &= \text{Concat}(u,\, E_i) \\ D(y) &= \text{Dropout}\left(\text{ReLU}(\text{BN}(\text{Conv}_{1\text{D}}^{k=3,p=1}(y)))\right) \\ D_{\text{out}}(y) &= \text{Dropout}(\text{ReLU}(\text{BN}\left(\text{Conv}_{1\text{D}}^{k=5,p=2}(D(y))\right))) \end{aligned} \tag{8}$$

The final stage applies a 1×1 convolution to reduce the channel dimension to d_{cond} (the number of geometric parameters), and an adaptive global pooling (AdaptiveAvgPool1d output size 1) collapses the temporal axis to produce the final prediction $\hat{P} \in \mathbb{R}^{d_{\text{cond}}}$.

Notation: B denotes the batch size; C is the channel count; L is the temporal length (frequency bins); and other symbols follow the conventions introduced for the forward model.

3. APPLICATION CASE OF PROPOSED METHODS

3.1. Antenna Case and Training Strategies

The UWB stepped microstrip monopole antenna [25], shown in Fig. 3, is employed as the case of the proposed methods. The antenna's dimensions are 13 mm × 22 mm × 0.787 mm, and it is fed by a stepped microstrip line. The radiation patch is a rectangular structure with two triangular missing slots, and the defected ground structure (DGS) is employed to optimize the performance of the antenna. The comprehensive physical parameters are provided in Table 1.

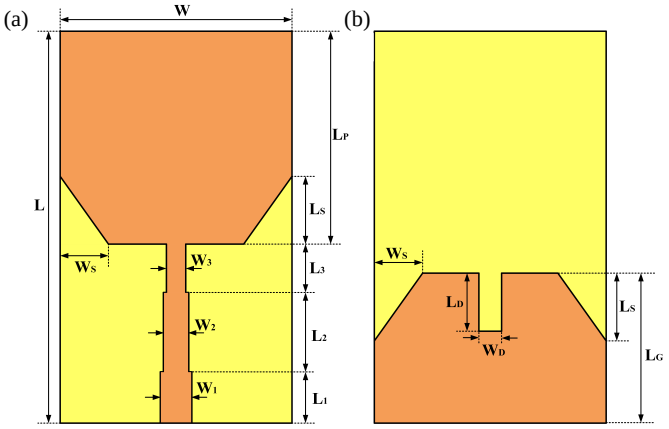

FIGURE 3. Structure of the antenna case: (a) top structure; (b) bottom structure.

TABLE 1. Physical parameters of the antenna case.

Parameter	Value (mm)	Parameter	Value (mm)
L	22	W	13
H	0.787	W_1	1.5
W_2	1.3	W_3	1
L_1	3	L_2	5
L_3	2	L_P	12
L_S	4.2	W_S	3.5
L_G	8.7	L_D	3.6
W_D	1.3		

Utilizing High Frequency Structure Simulator (HFSS) for the full-parameter scanning simulation, the S_{11} curve of the antenna is depicted in Fig. 4. The simulation outcomes demonstrate that the proposed antenna complies with the operational requirements within the frequency band spanning from 2.8 GHz to 22.3 GHz.

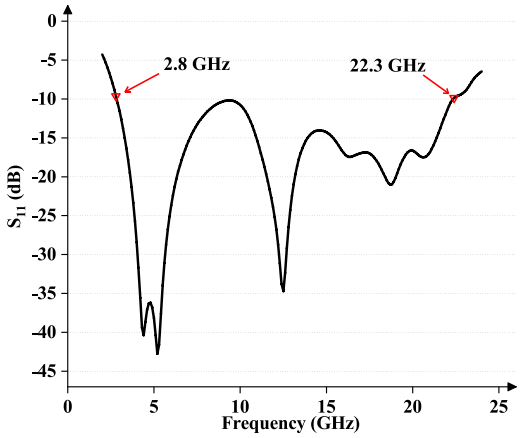

FIGURE 4. S_{11} curve of the antenna case simulated from HFSS.

The eight key physical parameters are denoted as $G = [L_D, L_G, L_S, W_1, W_2, W_3, W_D, W_S]$, and the corresponding value range and sampling step are set for each parameter, as shown in Table 2.

TABLE 2. Sampling ranges and steps of physical parameters.

Geometric parameter	Sampling range (mm)	Sampling step (mm)	Number of samples
L_D	3.6–3.7	0.05	3
L_G	8.6–8.7	0.05	3
L_S	4.1–4.2	0.1	2
W_1	1.4–1.6	0.1	3
W_2	1.2–1.3	0.05	3
W_3	0.9–1	0.05	3
W_D	1.27–1.3	0.015	3
W_S	3.4–3.5	0.1	2

The original dataset is constructed through the simulation of HFSS. The process yields totally 2916 sets of data via full factorial design, each comprising eight physical parameters and their mapped S_{11} curve sampling sequences. The frequency range of the S_{11} curve is set from 2 GHz to 24 GHz, and the sampling step size is set as 0.1 GHz, resulting in each S_{11} curve sampling sequence containing 221 values.

The dataset is divided into training, validation, and test sets according to the ratio of $7:1:1$, and the maximum training epochs of the two models are both set to 500, to ensure that the models can thoroughly learn complex patterns in the data.

During the training session, several training strategies are employed to enhance the accuracy and generalization of the models.

Firstly, the original data is normalized via min-max normalization strategy. Both physical parameters and S_{11} sequences are normalized into $[0, 1]$, respectively.

Secondly, the gradient updates in each stage through an AdamW optimizer [34], an extension of the Adam optimizer that decouples weight-decay from the adaptive moment estimation, which helps prevent the "vanishing" of the weight-decay penalty in the presence of per-parameter learning rates, yielding better generalization.

Moreover, learning rate scheduling strategies are different between the forward design and inverse design. In the forward design, set the initial learning rate (lr) of 1e-4, the ReduceLROnPlateau strategy with the patience (20 epochs) whenever the validation loss fails to decrease.

The lr decay factor (0.1) of the ReduceLROnPlateau strategy is introduced to dynamically decay the lr, which can maintain a high lr while the model is still improving, then sharply reduce it once progress stalls, thus striking a good balance between rapid initial convergence and fine-tuning stability.

While setting the initial lr (1e-3) in the inverse design, the maximum iterations ($T_{\max}$) of the CosineAnnealingLR strategy is 100 epochs which means that over the first 100 epochs, the lr undergoes a cosine-shaped decay from its initial value (1e-3) down to a very low value and that the remaining 400 epochs are then devoted to fine-tuning at this minimal learning rate.

The CosineAnnealingLR strategy ensures that the model can rapidly converge toward a promising region of the solution space during the early phase with relatively large update steps, while subsequently conducting a more precise search for improved optima with very small steps, thereby striking a balance between training efficiency and ultimate model accuracy.

Additionally, SmoothL1Loss is adopted as the loss function in each training process to reduce the influence of outliers while ensuring the convergence speed, thereby improving the stability and robustness of the model. SmoothL1Loss transits from an L2 Loss (Mean Squared Error, MSE) near zero to an L1 Loss (Mean Absolute Error, MAE) for larger residuals, achieving both sensitivity to small errors and robustness to occasional large outliers.

The experimental setup involves a computing device equipped with an Intel i7-13700H (@2.4 GHz) CPU, an Nvidia RTX 4060 GPU, and 32 GB memory. The two models are developed using the PyTorch 2.1.0 framework, and the programming language employed is Python 3.9.

3.2. Results Analyses of Proposed Forward Model

As the surrogate model of the traditional methods, accuracy and generalization are important for the forward model. Fig. 5 shows the training loss and validation loss during 500 epochs. From the figure, the two losses plummet to 0.02 fast and simultaneously in the first 50 epochs. Before 100 epochs, the two losses drop more gently to 0.02. After 100 epochs, the two

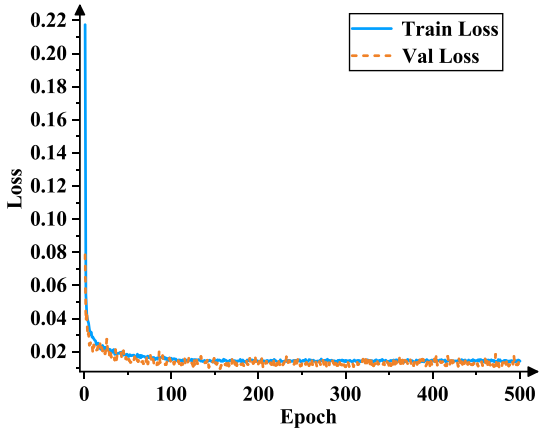

FIGURE 5. Training loss and validation loss of 1D ConDiffusion-U-Net.

losses fluctuate at 0.014 slightly, which indicates that the forward model has reached the convergence state.

To demonstrate improved predictive accuracy of the proposed forward model, DMLP [10] and 1D-MCNN [11] from our team lab are chosen as baseline models to conduct comparative analyses. Root mean square error (RMSE) and mean absolute error (MAE) are employed as metrics for evaluating the model prediction performance and are shown in Table 3.

TABLE 3. Performance comparison of three forward models.

Forward Model	RMSE	MAE
1D ConDiffusion-U-Net	0.117	0.0800
1D-MCNN	0.323	0.267
DMLP	0.395	0.286

From the table, 1D ConDiffusion-U-Net achieves markedly superior predictive accuracy on the test set compared to all the baseline models. Specifically, it reduces the RMSE to 0.117, about 63.8% and 70.4% decreasing against 1D-MCNN (RMSE $=$ 0.323) and DMLP (RMSE $=$ 0.395), and cuts the MAE to 0.0800, a 70.0% and 72.0% decreasing against 1D-MCNN (MAE $=$ 0.267) and DMLP (MAE $=$ 0.286).

These results demonstrate that conditioning a denoising diffusion process on antenna physical parameters enables the network to capture the complex, multi-scale mapping from design parameters to the S_{11} curve far more faithfully than the former proposed models. In practice, this translates into both sharper curve reconstructions and more accurate average predictions across the entire frequency spectrum. Moreover, the time cost is much lower than traditional methods: for 500 epochs of training, it is about 8 minutes, and for testing it is only 0.08 seconds per sample.

Randomly choosing one group of data not included in the dataset to validate the generalization performance of the proposed forward model, the original S_{11} curve from HFSS and the predicted S_{11} curves of various models are shown in Fig. 6.

From Fig. 6, the proposed forward model proves its better generalization performance than other forward models. In the low-frequency range close to 5 GHz, three models all predicted the curve with features of two minimum points, but the predicted curve of 1D ConDiffusion-U-Net is much closer to the original curve of HFSS than 1D-MCNN and DMLP.

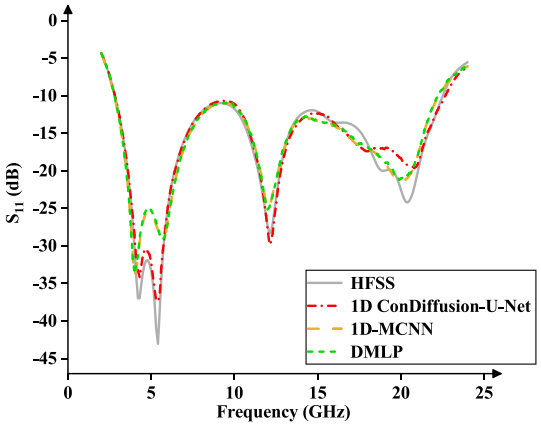

FIGURE 6. Original and predicted S_{11} curves from HFSS and forward models (Original S_{11} curve from HFSS and predicted S_{11} curves from 1D ConDiffusion-U-Net, 1D-MCNN and DMLP).

In the middle-frequency range close to 12 GHz, obviously, 1D ConDiffusion-U-Net shows the best-predicting performance that 1D ConDiffusion-U-Net obtains the lowest S_{11} minimum value which is almost equal to that of HFSS.

Because of the unstable changing trend of S_{11} values in the high-frequency range from 17 GHz to 23 GHz in the dataset, it is a great challenge for the forward model to capture highly accurate features in this range. Therefore, there are some deviations between the predicted curves and original curve. Although the predicted curve of 1D ConDiffusion-U-Net is not as close as the other two models to the original curve, more complicated changing features of more than one minimum point of the curve are predicted by 1D ConDiffusion-U-Net, thus remaining competitive with other models.

3.3. Results Analyses of Proposed Inverse Model

Figure 7 displays the training loss and validation loss of the inverse model, 1D U-Net-CSA. Both training loss and validation loss plummet before 10 epochs. As the training continues, both losses begin to gently decrease, and after 500 epochs, both losses are lower than 1.5e-5, which indicates that the inverse model reaches the convergence state with low errors.

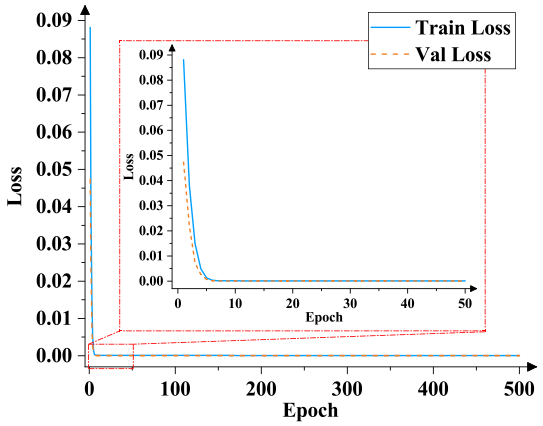

FIGURE 7. Training loss and validation loss of 1D U-Net-CSA.

Moreover, to verify the initial performance of the proposed model, 1D U-Net and MLP are set as baseline models. Fix the training epochs as 500 for three models, and set the RMSE and MAE as quantitative indexes, which are shown in Table 4.

From the table, the proposed inverse model clearly demonstrates the best overall performance. 1D U-Net-CSA achieves the lowest RMSE (0.0321), reducing the squared-error penalty

TABLE 4. Performance comparison of three inverse models.

Forward Model	RMSE	MAE
1D U-Net-CSA	0.0321	0.0226
1D U-Net	0.0409	0.0301
MLP	0.0521	0.0375

by roughly 21.5% compared to the standard 1D U-Net and by 38.4% versus the MLP. Its MAE (0.0226) also remains better than the standard MLP (0.0375) and the standard 1D U-Net (0.0301) with respective decreases of roughly 24.9% and 39.7%, indicating that 1D U-Net-CSA not only tames large outliers but also maintains uniformly small absolute errors.

Together, these results confirm that incorporating the attention mechanisms into the U-Net backbone yields the inverse model that is both more robust to extreme errors and highly accurate on average, making it a better choice for inverse design. Moreover, compared to the traditional methods, the proposed inverse model has much lower time costs: for 500 epochs of training, it is less than 200 seconds, and for testing it is only 0.3 seconds per sample.

One desired S_{11} curve is simulated which is not in the dataset via HFSS as the input of 1D U-Net-CSA. The predicted physical parameters are then sent back to HFSS to obtain the output S_{11} curve, and the comparison between the input S_{11} curve and output one is shown in Fig. 8.

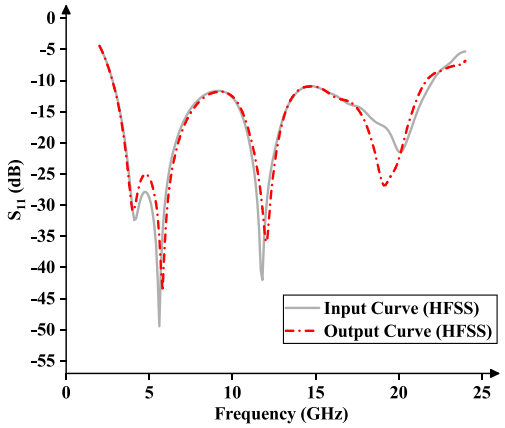

FIGURE 8. Input S_{11} curve of 1D U-Net-CSA and output S_{11} curve of the predicted physical parameters.

From the figure, the output S_{11} curve and input S_{11} curve show a high degree of agreement almost over the full frequency band, which proves the design validation of the proposed inverse model. Moreover, the output curve of the high-frequency range close to 20 GHz is much lower than that of the input one, indicating that the antenna with the predicted physical parameters can obtain a better matching performance. The outcomes of the figure demonstrate that the proposed inverse model not only can reach the design S_{11} target but also has extrapolation capabilities which implies that the inverse model has the potential to optimize the performance of the antenna.

The proposed two frameworks are not only predefined architectures but also formalized as conditional operators: the forward surrogate models the conditional density $p(x_0|P)$ via a diffusion-based reverse sampler (Eqs. (1)–(3)), with FiLM-conditioned U-Net noise predictors $\epsilon_\theta(x_t, t, P)$; the inverse surrogate defines a direct operator $\hat{P} = g_\phi(X_0)$ implemented by a U-Net augmented with channel and spatial attention (Eqs. (5)–(8)). Empirically, these mathematical design choices produce substantial improvements over baselines (forward: $\geq$ 63% RMSE and $\geq$ 70% MAE improvements; inverse: $\geq$ 21% RMSE and $\geq$ 38% MAE improvements). Structurally, diffusion conditioning, FiLM modulation, and multi-scale attention increase expressivity and enable modelling conditional distributions and scale-dependent features that simple MLP/MCNN baselines cannot capture.

4. CONCLUSION

To improve the design efficiency of antennas, this study has proposed two deep-neural-network-based antenna design methods which mainly focus on the relation between physical parameters and S_{11} curves. Both structures of the models fused various mechanisms including diffusion and attention mechanisms to improve the models' accuracy and efficiency. The forward model — 1D ConDiffusion-U-Net can accurately and efficiently predict S_{11} curves of the target physical parameters; the inverse model — 1D U-Net-CSA can directly and validly output the physical parameters of the target S_{11} curves for the antenna inverse design. The experimental outcomes through a UWB antenna case have proved the efficacy of the proposed methods. Moreover, more studies will be set in the future, including the nonuniqueness mitigation of the inverse model, multi-objectives design and optimization of antennas, the joint training of the forward model and inverse model to construct an integral and cyclic design-validation system, mitigating the potential interference factors arising during the fabricating process when using neural networks for antenna design.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China General Program under Grant 61971210.

REFERENCES

[1] Andriulli, F., P.-Y. Chen, D. Erricolo, and J.-M. Jin, "Guest editorial machine learning in antenna design, modeling, and measurements," *IEEE Transactions on Antennas and Propagation*, Vol. 70, No. 7, 4948–4952, 2022.

[2] Zhu, Z., Y. Tian, and J. Sun, "Antenna modeling based on image-CNN-LSTM," *IEEE Antennas and Wireless Propagation Letters*, Vol. 23, No. 9, 2738–2742, 2024.

[3] Pantanowitz, L., T. Pearce, I. Abukhiran, M. Hanna, S. Wheeler, T. R. Soong, A. P. Tafti, J. Pantanowitz, M. Y. Lu, F. Mahmood, *et al.*, "Nongenerative artificial intelligence in medicine: Advancements and applications in supervised and unsupervised machine learning," *Modern Pathology*, Vol. 38, No. 3, 100680, 2025.

[4] Kuznietsov, A., B. Gyevnar, C. Wang, S. Peters, and S. V. Albrecht, "Explainable AI for safe and trustworthy autonomous driving: A systematic review," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 25, No. 12, 19 342–19 364, 2024.

[5] Xiong, G., P. Helo, S. Ekström, and Z. Shen, "A service-oriented autonomous crane system," *IEEE Transactions on Computational Social Systems*, Vol. 11, No. 6, 8030–8045, 2024.

[6] Koziel, S., N. Çalık, P. Mahouti, and M. A. Belen, "Accurate modeling of antenna structures by means of domain confinement and pyramidal deep neural networks," *IEEE Transactions on Antennas and Propagation*, Vol. 70, No. 3, 2174–2188, 2022.

[7] Koziel, S., N. Çalık, P. Mahouti, and M. A. Belen, "Low-cost and highly accurate behavioral modeling of antenna structures by means of knowledge-based domain-constrained deep learning surrogates," *IEEE Transactions on Antennas and Propagation*, Vol. 71, No. 1, 105–118, 2023.

[8] Wei, Z., Z. Zhou, P. Wang, J. Ren, Y. Yin, G. F. Pedersen, and M. Shen, "Fast and automatic parametric model construction of antenna structures using CNN-LSTM networks," *IEEE Transactions on Antennas and Propagation*, Vol. 72, No. 2, 1319–1328, 2024.

[9] Su, Y., Y. Yin, H. Chen, S. Li, H. Zhao, Z. Su, and X. Yin, "Time-domain scattering parameters-based neural network inverse model for antenna designs," *IEEE Antennas and Wireless Propagation Letters*, Vol. 23, No. 7, 1976–1980, 2024.

[10] Nan, J., Y. Du, M. Wang, and M. Gao, "Deep learning architecture and neural network optimization of ultra-wideband antenna modeling," *Laser & Optoelectronics Progress*, Vol. 59, No. 13, 1323001, 2022.

[11] Nan, J., W. Sun, Y. Du, and M. Wang, "One-dimensional convolutional neural network modeling method for ultra-wideband antenna," *Journal of Electronic Measurement and Instrument*, Vol. 37, No. 2, 204–210, 2023.

[12] Peng, F. and X. Chen, "A low-cost optimization method for 2D antennas using a disassemblable convolutional neural network," *IEEE Transactions on Antennas and Propagation*, Vol. 72, No. 9, 7057–7067, 2024.

[13] Huang, H., X.-S. Yang, and L. Yuan, "Antenna shape neural network modeling based on computer vision," in *National Conference on Microwave and Millimeter Wave (NCMMW'21)*, 691–693, Nanjing, China, May 2021.

[14] Ninković, D., S. B. Shah, A. Altunaiji, N. Ali, and D. Olćan, "Comparison of ensembles of deep neural networks and mixture of experts for antenna modeling," *IEEE Antennas and Wireless Propagation Letters*, 1–5, 2025.

[15] Liu, P., D. Liu, Y. Li, S. Ye, and D. Su, "Attention-based resnet for radiation pattern prediction of phased array antenna," *IEEE Antennas and Wireless Propagation Letters*, Vol. 23, No. 12, 4453–4457, 2024.

[16] Jin, J., Q. Su, Y. Xu, Z. He, and Y. Lu, "Efficient radiation pattern prediction of array antennas based on complex-valued graph neural networks," *IEEE Antennas and Wireless Propagation Letters*, Vol. 21, No. 12, 2467–2471, 2022.

[17] Karahan, E. A., A. Gupta, U. K. Khankhoje, and K. Sengupta, "Deep learning based modeling and inverse design for arbitrary planar antenna structures at RF and millimeter-wave," in *2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI)*, 499–500, Denver, CO, USA, Jul. 2022.

[18] Gosal, G., D. A. McNamara, and M. C. E. Yagoub, "The use of inverse neural networks in transmitarray antenna design," in *2014 IEEE Antennas and Propagation Society International Symposium (APSURSI)*, 1272–1273, Memphis, TN, USA, 2014.

[19] Liu, J.-P., B.-Z. Wang, C.-S. Chen, and R. Wang, "Inverse design method for horn antennas based on knowledge-embedded physics-informed neural networks," *IEEE Antennas and Wireless Propagation Letters*, Vol. 23, No. 6, 1665–1669, 2024.

[20] Shereen, M. K., X. Liu, X. Wu, A. Naseem, and M. Uzair, "Deep learning-inspired linear regression technique for accurate microstrip antenna performance analysis," in *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, 42–47, Yogyakarta, Indonesia, 2025.

[21] Yao, H. M., M. Li, L. Jiang, K. L. Yeung, and M. Ng, "Antenna array diagnosis using a deep learning approach," *IEEE Transactions on Antennas and Propagation*, Vol. 72, No. 6, 5396–5401, 2024.

[22] Chen, R., C. W. L. Lee, P. K. Tan, and T. H. Gan, "Reflectarray antenna design using the deep learning controlnet diffusion model," in *2025 19th European Conference on Antennas and Propagation (EuCAP)*, 1–5, Stockholm, Sweden, 2025.

[23] Nan, J., X. Cao, M. Gao, and P. Zhang, "A research method for reverse modeling of ultra-wideband antenna with dual notch characteristics based on ALO-LMBP neural network," in *National Conference on Microwave and Millimeter Wave (NCMMW'21)*, 354–356, Nanjing, China, May 2021.

[24] Nan, J. and M. Wang, "Design and research of ultra-wideband stepped microstrip monopole antenna," *Chinese Journal of Radio Science*, Vol. 36, No. 2, 225–230, 2021.

[25] Ronneberger, O., P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241, Munich, Germany, Oct. 2015.

[26] Tatli, U. and C. Budak, "Biomedical image segmentation with modified U-Net," *Traitement du Signal*, Vol. 40, No. 2, 523–531, 2023.

[27] Senapati, P., A. Basu, M. Deb, and K. G. Dhal, "Sharp dense U-Net: An enhanced dense U-Net architecture for nucleus segmentation," *International Journal of Machine Learning and Cybernetics*, Vol. 15, No. 6, 2079–2094, 2024.

[28] Xu, J., D. Zeng, and J. Xi, "High resolution remote sensing semantic segmentation using bayesian of hyperparameters and improved U-Net," in *2023 China Automation Congress (CAC)*, 7025–7029, Chongqing, China, Nov. 2023.

[29] Rajamani, K. T., P. Rani, H. Siebert, R. E. Ramalingam, and M. P. Heinrich, "Attention-augmented U-Net (AA-U-Net) for semantic segmentation," *Signal, Image and Video Processing*, Vol. 17, No. 4, 981–989, 2023.

[30] He, K., X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European Conference on Computer Vision (ECCV'16)*, 630–645, Amsterdam, The Netherlands, Oct. 2016.

[31] Perez, E., F. Strub, H. d. Vries, V. Dumoulin, and A. Courville, "FiLM: Visual reasoning with a general conditioning layer," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'18)*, 3942–3951, New Orleans, USA, Apr. 2018.

[32] Hu, J., L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *IEEE Computer Society*, Vol. 42, No. 8, 2011–2023, 2020.

[33] Woo, S., J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV'18)*, 3–19, Munich, Germany, Sep. 2018.

[34] Loshchilov, I. and F. Hutter, "Decoupled weight decay regularization," in *International Conference on Learning Representations (ICLR'19)*, 1–8, New Orleans, USA, May 2019.