International Journal of Microwave and Wireless Technologies

cambridge.org/mrf

A reinforcement learning-driven automated antenna topology optimization method by using enhanced hybrid multilayer perceptron

Jiangling Dou[1,2], Chenyang Zhao[2], Jian Song[2] and Tao Shen[3]

[1]Yunnan Key Laboratory of Computer Technologies Application, Kunming, China; [2]Information Engineering and Automation,Kunming University of Science and Technology, Kunming, China and [3]Yunnan Electromechanical Vocational and Technical College, Kunming, China

Research Paper

Cite this article: Dou J, Zhao C, Song J, Shen T (2025) A reinforcement learning-driven automated antenna topology optimization method by using enhanced hybrid multilayer perceptron. *International Journal of Microwave and Wireless Technologies* **17**(10), 1841–1850. https://doi.org/10.1017/S1759078725102559

Received: 25 June 2025
Revised: 19 October 2025
Accepted: 28 October 2025

Keywords:
antenna topology optimization; enhanced hybrid multilayer perceptron (E-MLPs); microstrip antenna; reinforcement learning (RL)

Corresponding author:
Tao Shen;
Email: shentao@kust.edu.cn

Abstract

A reinforcement learning (RL)-based automated antenna topology optimization method is proposed. The proposed framework can be divided into three phases, which are high-quality dataset construction, electromagnetic (EM) simulation acceleration, and RL-driven automated antenna topology optimization. Based on the high-quality dataset, a fully trained enhanced hybrid multilayer perceptron is proposed to replace time-consuming EM simulations. This approach allows the RL to acquire knowledge from the interaction between antenna topology and the environment quickly, reducing the optimization time cost caused by the large number of EM simulations. Additionally, two crucial components, topology bidirectional mapping strategy (TBM) and topology hierarchical analyzation strategy (THA), are introduced in this work to address the compatibility problems between ML and high-dimension antenna topology data. To verify the effectiveness of the proposed method, a microstrip patch antenna operating at 2.45 GHz is optimized. According to the measurement results, the antenna performance of gain and impedance bandwidth is improved greatly at the same time through the proposed method.

Introduction

With the rapid development of artificial intelligence technology, the antenna optimization methods assisted by machine learning (ML) have provided designers an efficient antenna design solution. Various ML techniques, including support vector machine (SVM) [1], multilayer perceptron (MLP) [2], Gaussian process regression (GPR) [3], and artificial neural network (ANN) [4], have been widely used in the field of electromagnetic (EM) design. By constructing a nonlinear mapping relationship between antenna geometric parameters and EM responses, the antenna optimization efficiency can be significantly improved. Furthermore, researchers have attempted to combine evolutionary algorithms (EAs), such as particle swarm optimization (PSO) [5], genetic algorithm (GA) [6] and grey wolf optimization algorithm (GWO) [7], with ML to establish a collaborative optimization method of ML-assisted evolutionary algorithms (ML-EAs) in recent years. For example, a hybrid ANN-SA optimization framework was designed in [8], which significantly improved the impedance bandwidth of a patch antenna. In [9], researchers had proposed a method that combined quantum genetic algorithm (QGA) with fully connected neural network to optimize the impedance bandwidth and gain of a stacked antenna array.

However, although the antenna performance can be enhanced by fine-tuning the initial parameters, the optimization effectiveness of the traditional geometric parameter optimization method is still hindered by the antenna predefined structural templates [10]. In order to solve this problem, researchers have been devoted to study the antenna topology optimization method based on ML. In [11–13], different inverse ML models were constructed to map the performance indicators to antenna topology. This method allows to generate different antenna structures that meet the performance requirements reversely. However, to ensure the accuracy of the inverse ML model, a large amount of data will be required for the training. When the input performance deviates from the knowledge learned by the model, the antenna performance predictive ability will be deteriorated, resulting in the inability to provide an accurate design solution. In addition, the overall optimization effect of ML will be also seriously restricted by the high-dimension antenna topology data.

Therefore, a novel framework for antenna topology optimization method is introduced in this paper. It contains three core modules, which are the dataset preparation module driven by QGA, the agent prediction module based on enhanced hybrid multilayer perceptron (E-MLPs), and the decision optimization module based on reinforcement learning (RL). As the main modules of the proposed method, E-MLPs and RL are responsible for the agent prediction

© The Author(s), 2026. Published by Cambridge University Press in association with The European Microwave Association

and decision optimization task of topology design, respectively. Additionally, to address the algorithm compatibility problem with antenna topology data, topology bidirectional mapping strategy (TBM) and topology hierarchical analyzation strategy (THA) are proposed. TBM establishes a mathematical equivalence relationship between antenna physical topology and binarized matrix. THA restructures antenna topology into multi-channel data structure, which fully utilize the cross-channel data mixing ability of E-MLPs. Finally, a microstrip patch antenna that operating at 2.45 GHz is optimized by the proposed method. Compared with the initial design, the working frequency of antenna impedance bandwidth is increased from 2.40 GHz-2.43 GHz to 2.43 GHz-2.50 GHz, and the gain is increased from 2.55 dBi to 2.90 dBi in this work.

Proposed method

The overall implementation framework of the proposed method

The framework of the design method is shown in Fig. 1 QGA is initially employed as a data generator to collect potential high-quality data continuously. The topology data obtained by QGA is then converted into the antenna model and sent to HFSS for simulation. Subsequently, E-MLPs, which is used to accelerate the EM simulation, is trained by the dataset with the proposed method. During the optimization, the training dataset for E-MLPs will be continued to collect. Finally, RL is used to optimize the antenna topology. RL can dynamically adjust each decision based on the rewards associated with different actions in the design environment [14, 15]. It is worth noting that QGA is only applied in the initial dataset collected stage rather than the final antenna optimized stage. The primary reason is that EAs requires multiple full-wave EM simulations to get the optimal topology solution [16–18].

The method of training dataset construction and processing

In this paper, a high-quality dataset construction method is proposed and its implementation framework is shown in Fig. 2. Compared with traditional orthogonal or random sampling method, the proposed method ensures the quality of the dataset. In addition, during the process of dataset collection, TBM is developed to define the bidirectional conversion between the antenna topology structure and the antenna topology coding matrix, which is aimed to assist the algorithm to identify the antenna topology data. As shown in Fig. 2, the initial antenna topology is decomposed into several basic topology grid units. Then the antenna topology can be reconstructed by adjusting the material distribution of each basic topology grid unit. The antenna topology coding matrix can be expressed as

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}_{m\times n}, \quad x_{i,j} \in \{0,1\} \tag{1}$$

where n and m represent the number of antenna horizontal and vertical topology grid units, respectively. The binary value $x_{i,j} \in \{0,1\}$ in the matrix X indicates the material coverage state of each grid unit. Specifically, the grid at (i,j) of the antenna has the metal material coverage only when $x_{i,j} = 1$. Each possible antenna topology is stored in the population of QGA, which is composed of several binary value $x_{i,j}$.

Finally, the collected data is processed by using THA. The initial topology matrix is firstly divided into independent components and then augmented along the additional dimension. Through THA, the 2D antenna topology matrix is effectively transformed into a three-channel tensor structure X', which is similar to the framework that commonly used in computer vision.

The proposed E-MLPs architecture

In this paper, the proposed E-MLPs adopts the main architecture of MLP-Mixer, which is originally designed for computer vision tasks [19]. The special feature of E-MLPs lies in its MLP-based information mixing mechanism. This mechanism avoids expensive convolution or attention-based mechanisms. It efficiently captures global relationships across spatial and feature dimensions [20, 21]. In this work, the proposed E-MLPs is introduced to the field of EM design innovatively, which incorporates three key design modifications as follows.

1) Use the 3D topology tensor as the input data of the model. This special 3D antenna topology tensor can fully utilized the cross-channel information mixing capability of E-MLPs, allowing the topology characteristics of the antenna to be comprehensively captured.

2) Establish a nonlinear mapping from antenna topology to EM response by E-MLPs. The input of E-MLPs is restructured into antenna topology tensors and the EM indicators, such as impedance bandwidth and gain, are obtained from the output layer directly.

3) Build an online optimization mechanism for E-MLPs during the optimization process. In order to enhance the accuracy of the model, E-MLPs is improved with the online learning mechanism. Specifically, the top n best candidates in each iteration are selected for real EM simulation. The simulated results are then added to the dataset to retrain the E-MLPs.

The overall framework of the E-MLPs model is shown in Fig. 3, which consists of three parts. The input data is firstly divided into M complementary overlapping patches by the linear segmentation part and fed into the information mixing part which integrates token mixing layer and channel mixing layer. Token mixing layer establishes associations between different positions in the spatial dimension and integrates features through independent cross-channel operations. Channel-mixing layer realizes feature fusion in the channel dimension and performs cross-channel information interaction for each spatial position. Finally, the prediction and output part is used to fuse high-dimension feature tensors, establishing a direct mapping from the hybrid feature space to the antenna EM performance indicators. Notably, the information mixing layer is only implemented by the MLP, where each MLP comprises two fully connected layers. Each hidden layer contains 1024 neurons with GELU activation applied between the fully connected layers. The operations of token mixing and channel-mixing can be expressed as

$$Y^T = X_{*,i} + w_1 \,\mathrm{GELU}\left(w_2 \times \Theta(X'_{*,i})\right), \quad \text{for } i = 1, 2, \ldots, Q \tag{2}$$

$$Y^C = Y^T_{j,*} + w_3 \,\mathrm{GELU}\left(w_4 \times \Theta(Y^T_{j,*})\right), \quad \text{for } j = 1, 2, \ldots, M \tag{3}$$

where $\mathbf{Y}^T$ and $\mathbf{Y}^C$ represent the output results of token mixing layer and channel mixing layer, respectively. $\Theta(\cdot)$ is the normalization function, which is used to normalize the input and stabilize the neural network training process. GELU$(\cdot)$ is the activation function, which can be approximated as:

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

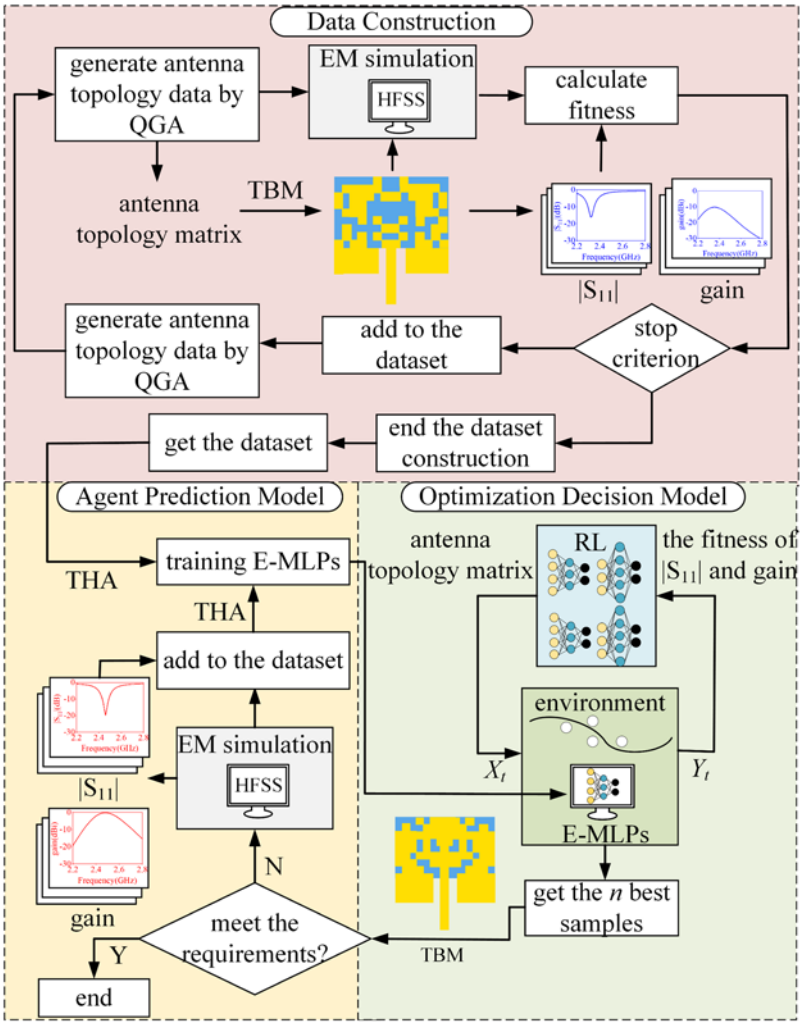


Figure 1. The framework of the proposed method.

$$\mathrm{GELU}(x) = \frac{1}{2}x\left(1 + \tanh\left(\sqrt{\frac{2}{\pi}}\left(x + 0.044715x^3\right)\right)\right) \quad (4)$$

During the training process, the loss between the predicted results and the real antenna performances is calculated through (5) to guide the process of forward propagation and update the network weights. The loss is defined as:

$$\mathrm{Loss}(y_i|\hat{y}_i) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\frac{y_{i,j} - \hat{y}_{i,j}}{y_{i,j}}\right)^2 \quad (5)$$

where N denotes the number of samples in each iteration during training. M denotes the number of antenna performance indicators. $y_{i,j}$ and $\hat{y}_{i,j}$ represent the real and prediction antenna performance indicators, such as impedance bandwidth and gain.

The proposed RL architecture

In this work, RL is proposed to optimize the antenna topology based on the deep deterministic policy gradient (DDPG) algorithm. DDPG is a model-free offline policy RL algorithm based on the actor-critic architecture [22]. The main framework of DDPG-based RL algorithm is shown in Fig. 4.

Actor network performs a certain mapping from the state s_t to the action a_t based on the policy μ. The original actor network parameters θ^μ are then updated using the gradient ascent method, which is determined by the gradient of the critic network and its own network. Critic network evaluates state-action pairs (s_t, a_t) to get the Q value function of the current policy and realizes policy evaluation by minimizing the temporal difference error (TD-loss). The parameter of the original critic network θ^Q is updated by calculating the gradient of TD-loss. In our work, RL is designed to optimize the policy of the algorithm by maximizing the Q-value function which is defined as the future expected discounted reward for a given state-action pair (s_t, a_t). Q-value function can be expressed as follows

$$Q(s_t, a_t) = \mathbb{E}\left[R_t \mid s_t, a_t\right] \quad (6)$$

where $\mathbb{E}[\cdot]$ denotes the expected value and R_t is the sum of discounted rewards achieved by learning the future steps, which can be calculated by (7).

$$R_t = \sum_{k=0}^{\infty}\gamma^k r_{t+k+1} \quad (7)$$

where the value of discount factor γ is between 0 and 1, which is used to adjust the relative importance of immediate rewards

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

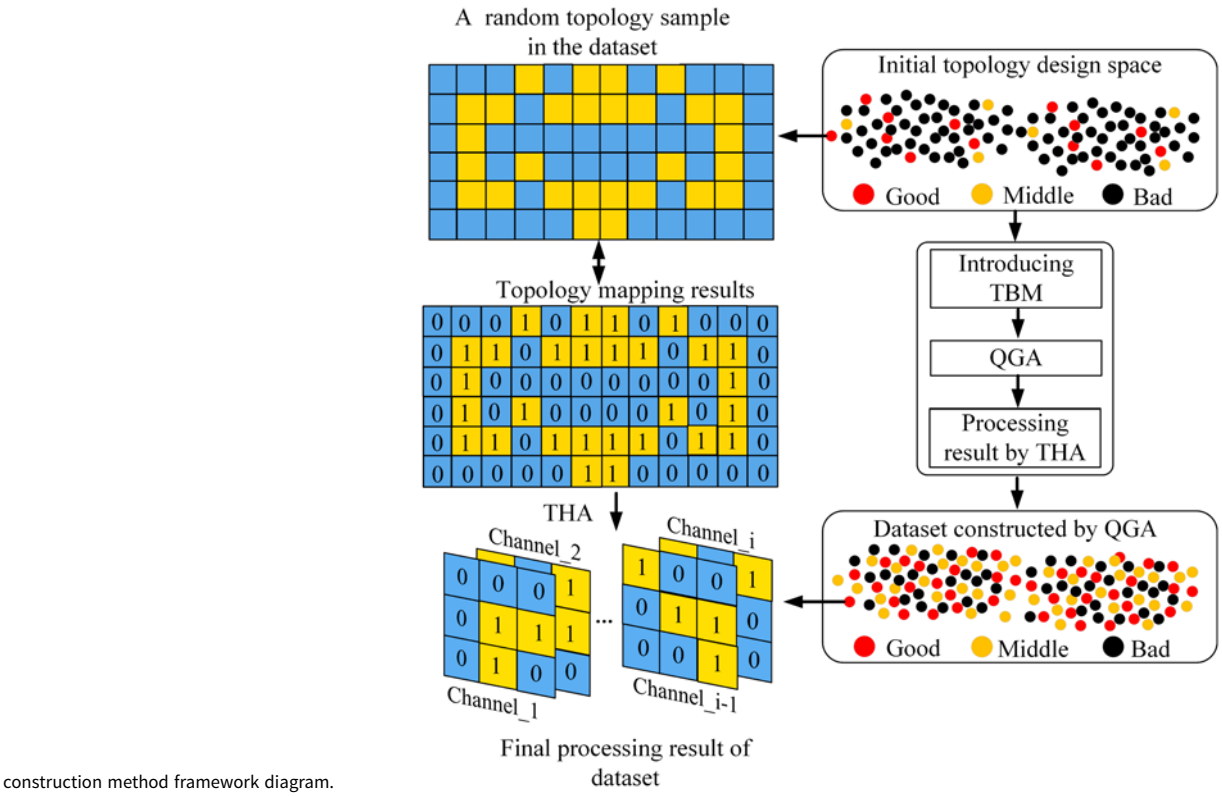


Figure 2. Dataset construction method framework diagram.

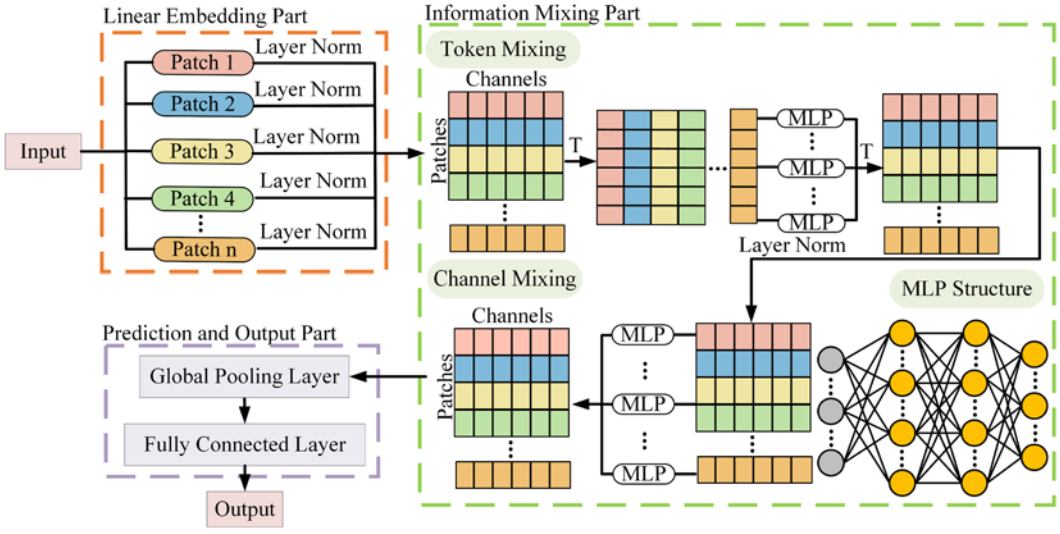


Figure 3. The framework diagram of E-MLPs.

and future rewards. In addition, to enhance the stability of the algorithm, two additional mechanisms, the target network and the experience replay buffer, are added to the DDPG. The target network can stabilize the training process by providing a fixed target for the network, with soft updates applied to both target networks. Through the soft update mechanism, the parameters of the original networks can be inherited by the target networks gradually. The experience replay buffer stores various experiences, including state s_t, action a_t, reward r_t, and next state s_{t+1}, which enables the interaction between RL and environment.

In this paper, the algorithm runs for i iterations and each round of iterations runs T steps. The key parameters of RL in the antenna optimization are as follows.

1) Action: Action a_t stores the topology information of the antenna. It can be expressed as (8)

$$a_{n,t} = (\alpha_{1,t}, \alpha_{2,t}, \alpha_{3,t}, \ldots, \alpha_{n,t}) \tag{8}$$

where $\alpha_{n,t}$ is the value of each topology grid unit which predicted by the actor network. The antenna topology X_t can be obtained by rearranging the action $\alpha_{n,t}$. $\alpha_{n,t}$ can be calculated by

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

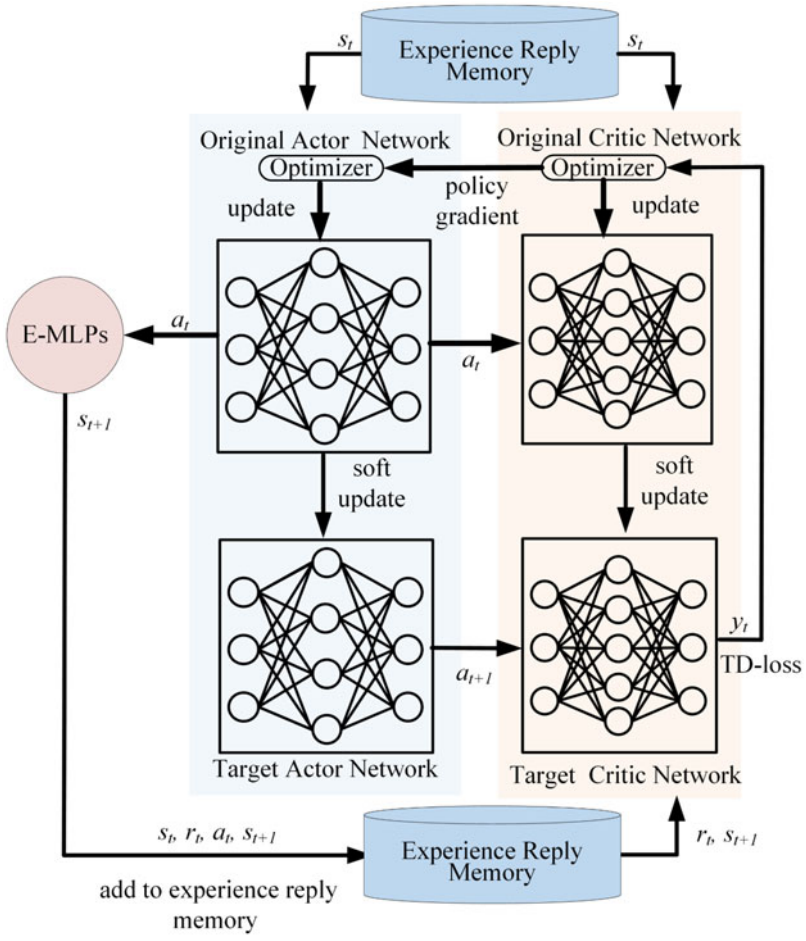


Figure 4. The framework diagram of RL.

$$\alpha_{n,t} = \begin{cases} 1, & \alpha_{n,t} \geq 0.5 \\ 0, & \alpha_{n,t} < 0.5 \end{cases} \tag{9}$$

2) State: The state s_t is defined as the performance of the antenna. In this work, in order to reduce the impact of high-dimension curve data on the prediction accuracy, the fitting value of the antenna performance Ψ_t is directly stored in the state s_t.

$$s_t = \Psi_t = \delta_1 \cdot \sum_{f=f_1}^{f_2} g_1(X_f, f) + \delta_2 \cdot g_2(f_0) \tag{10}$$

where $\delta_1 \epsilon (1, 0)$ and $\delta_2 \epsilon (1, 0)$ are the different personalized weights. They are used to describe the importance of gain and impedance bandwidth, aiming to help RL balance different performance indicators. In this paper, the value of δ_1 and δ_2 are set to 0.8 and 0.2, respectively. The $g_2(f_0)$ represents the gain value at the target frequency f_0. Frequency f_1 and f_2 are the limit of the target frequency band. In this paper, they are set to 2.30 GHz and 2.60 GHz, respectively. The $g_1(X_t, f)$ can be defined as

$$g_1(X_t, f) = \begin{cases} 1, & \text{if } |S_{11}(X_t, f)| \leq -10 \text{ dB} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where $S_{11}(X_t, f)$ represents the value of impedance matching when antenna topology X_t works at frequency f.

3) Reward: Reward r_t is a key parameter to update network. The reward function defines the immediate reward that calculated by a certain state. The main goal of the RL is to optimize the action strategy by maximizing the discounted cumulative reward. According to different design performance requirements, reward function should be given to ensure that the next action a_{t+1} can be correctly guided by RL. In this paper, the reward r_t is related to the fitting value Ψ_t, which is designed to be

$$r_t = \begin{cases} \dfrac{\Psi_t - \Psi_{t-1}}{\Psi_{\text{thre}}}, & \text{if } \Psi_t \geq \Psi_{\text{thre}} \text{and } \Psi_t \geq \Psi_{t-1} \\ 0, & \text{if } \Psi_{\text{thre}} < \Psi_t \leq \Psi_{t-1} \\ -0.1, & \text{if } \Psi_{t-1} \leq \Psi_t \leq \Psi_{\text{thre}} \\ -1, & \text{others} \end{cases} \tag{12}$$

where Ψ_{t-1} is the fitness value at the previous time step. Ψ_{thre} is the maximum threshold of the fitness value.

Analysis of experimental results

Introduction of antenna to be optimized

This section takes a miniaturized microstrip patch antenna as an example to verify the effectiveness of the proposed method, which is shown in Fig. 5. The antenna uses a single layer of FR4 epoxy resin as the dielectric substrate. The thickness of the substrate is 1.6 mm, and the relative dielectric constant is 4.4. The length of the patch W_p and the width of the patch L_p are 34 mm and 29 mm, respectively. The length of the ground W_s and the width of the grand L_s

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

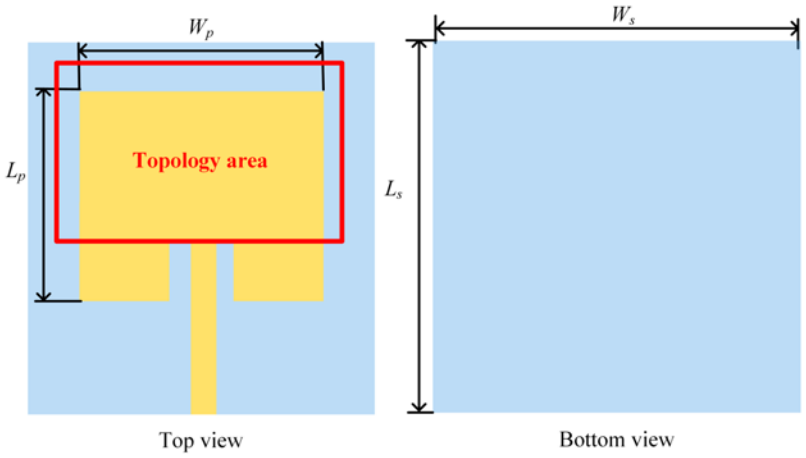


Figure 5. Initial antenna topology diagram.

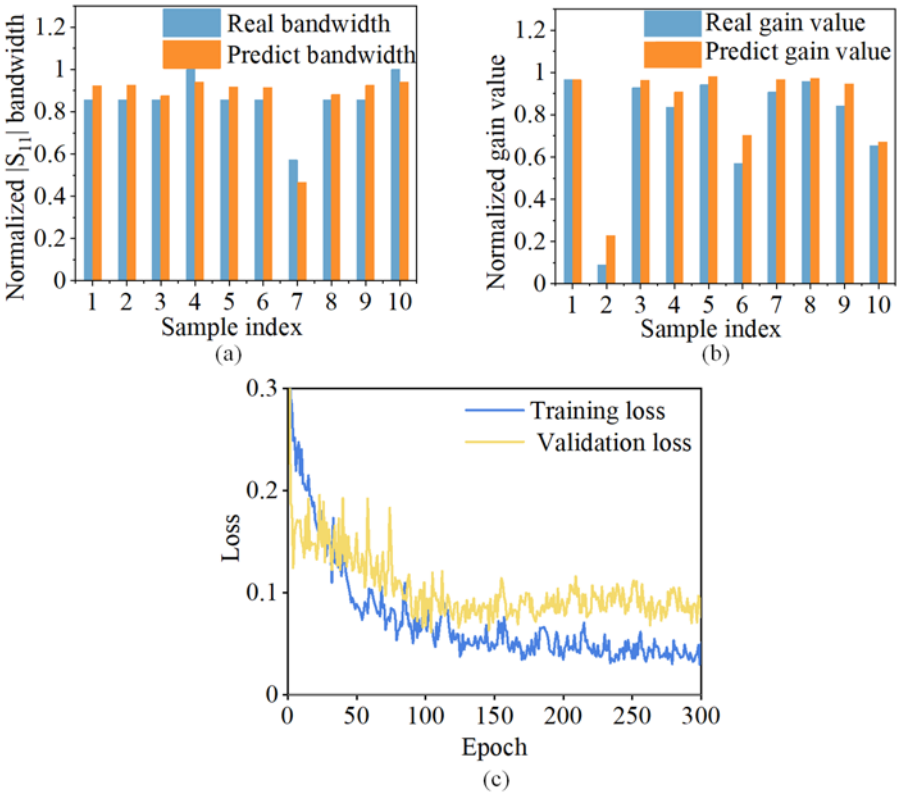


Figure 6. (a) The performance comparison of E-MLPs prediction results and 10 random samples for normalized $|S_{11}|$ bandwidth. (b) The performance comparison of E-MLPs prediction results and 10 random samples for normalized gain value. (c) The training loss and validation loss of the E-MLPs.

are 43 mm and 46 mm, respectively. In order to ensure the radiation characteristics of the antenna can be stimulated, the topology area is designed in the square part of the antenna.

Experimental environment and algorithm implementation

The entire simulation process is performed on a computer equipped with 128 GB memory and a 2.90 GHz Intel Xeon Platinum processor. The simulation software is ANSYS HFSS 2021R2. To realize conversion between the antenna topology and the antenna topology matrix that defined in TBM, the matrix can be converted into a CAD file and then imported into HFSS. It realizes the automated interaction between the algorithm and HFSS in this paper. Through the sampling method proposed in this paper, 1000 samples are collected for training. The number of design grid units of the antenna is set to 126, with the geometric symmetry constraints applied to the radiation patch. In addition, the original antenna topology is divided into 3 channels through the THA with each channel configured as a 3×7 grid unit. Finally, the trained E-MLPs is used to replace EM simulation for optimization. During the optimization, 100 iterations is performed in this work, each of iteration has 50 steps. In each iteration, the design results are obtained through E-MLPs and added to the experience reply memory. Then the 10 best candidates are selected for full-wave EM evaluation in the proposed method. The EM simulation results are stored in the dataset and used to retrain the E-MLPs. In addition, in order to speed up the optimization process, the candidate results with zero fitting values are skipped the EM simulation and are directly added to the dataset. This strategy effectively avoids unnecessary EM simulation during the early exploratory phase of RL.

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

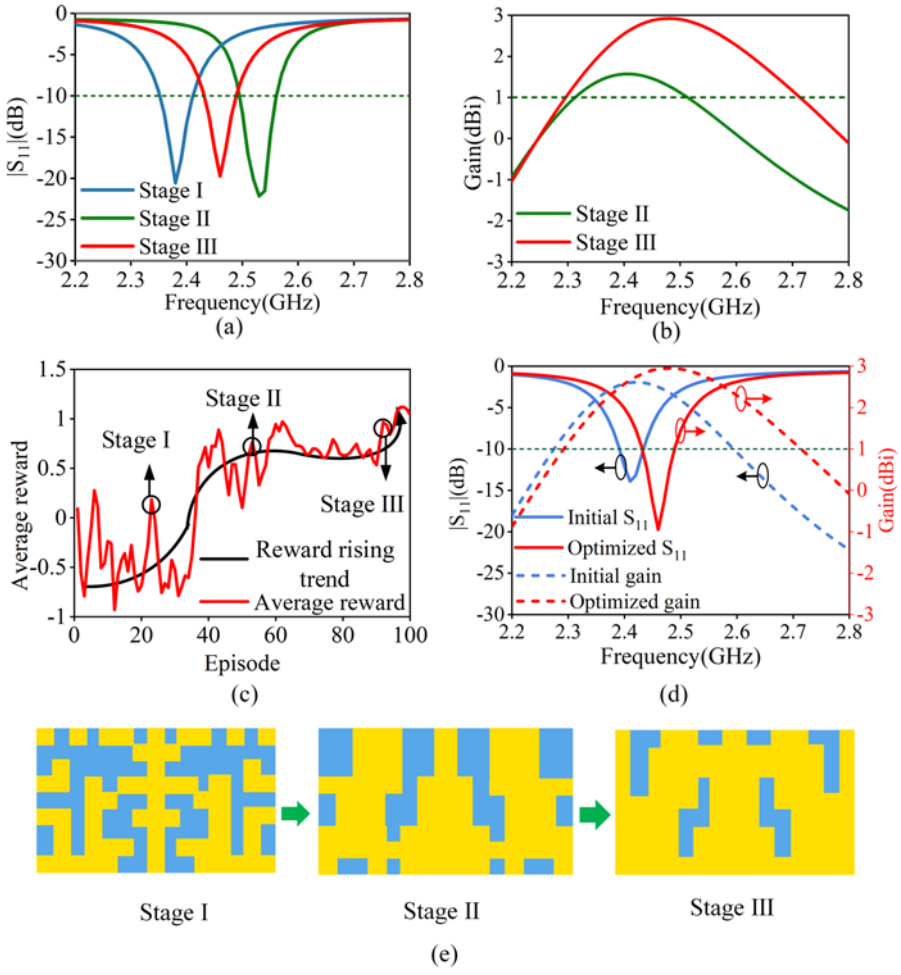


Figure 7. (a) The optimized results of $|S_{11}|$ in different stages. (b) The optimized results of gain in different stages. (c) The average reward of RL. (d) Final optimization results. (e) Antenna topology predicted by RL in different stages.

The loss curves for the training and validation results are shown in Fig. 6c. After 100 epochs full training, the training loss dropped around 0.03 and the validation loss remained around 0.1. Fig. 6a and 6b show the comparison between the predicted results and the actual antenna performances of 10 random datasets for the impedance bandwidths and gain values, respectively. It is important to note that both the gain values and impedance bandwidths have been normalized. The results indicate that although there are some variations between the predicted results and actual antenna performances, the overall predicted results can be accepted.

RL optimization results

In this paper, the requirements for antenna optimization are as follows:

1) Target #1: $|S_{11}| \leq -10$ dB from 2.40 GHz to 2.50 GHz.
2) Target #2: The operating bandwidth of antenna must include 2.45 GHz.
3) Target #3: The gain at 2.45 GHz must be greater than 1 dBi.

Fig. 7a and 7b show the antenna optimization results in the different optimization stages. Where the target threshold Ψ_{thre} is set to 5. For ease of presentation, three stages, the 1st iteration to the 30th iteration, the 31st iteration to the 60th iteration, and the 61st iteration to the 90th iteration, are referred to stage I, stage II, and stage III, respectively. In stage I, impedance bandwidth is successfully expended by RL. However, its operating frequency band is not the primary frequency band for 5G communication and its gain value at 2.45 GHz is lower than 0 dBi, failing to meet the minimum standard required for the engineering. Therefore, the gain curve of stage I is omitted in Fig. 7b. In stage II, the gain of antenna is significantly improved, but the frequency range when $|S_{11}| \leq -10$ still deviates from the target operating frequency. In stage III, the final performance optimization result is achieved by RL, resulting in significant improvement in both gain and impedance bandwidth.

The change of average reward is shown in Fig. 7c, which matches the RL optimization results of each stage. In stage I, the average reward is low and fluctuates significantly, which indicates that the algorithm is extensively exploring the topology design space. With continuous optimization, the average reward is improved in stage II but notable fluctuations remain, which can be explained by the fact that RL is still in the exploration of the topology space. Finally, the average reward gradually stable and reached the maximum value in stage III. The final optimization results can be also obtained. During the optimization, although the range of reward value during optimization is deeply affected by the design of the reward function, it has no impact on the ability of RL to maximize the cumulative rewards [23, 24].

The performance comparison between the initial antenna and the optimized antenna is shown in Fig. 7d. The frequency range of $|S_{11}| \leq -10$ is expanded from the initial 2.40 GHz-2.43 GHz to 2.43 GHz-2.50 GHz. Compared to the initial topology, the impedance bandwidth has been enhanced by more than 100%.

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

Table 1. Performance comparison between the initial and optimization results

Optimization stage	*$\|S_{11}\|$ bandwidth under -10 dB*	*$\|S_{11}\|$ at 2.45 GHz*	*Gain at 2.45 GHz*
Origin model	*2.40 GHz-2.43 GHz*	*−6.96 dB*	*2.55 dBi*
Stage I	*2.35 GHz-2.41 GHz*	*−4.65 dB*	*−18.84 dBi*
Stage II	*2.49 GHz-2.56 GHz*	*−2.92 dB*	*1.46 dBi*
Final results(Stage III)	*2.43 GHz-2.50 GHz*	*−15.62 dB*	*2.90 dBi*

Table 2. Comparison result with other works

Ref.	*Proposed method*	*Number of EM*	*Time*	*Object*	*Design space*
[25]	*GA*	8000	*Not given*	*$\|S_{11}\|$*	*198 grid units*
[26]	*HTBPSO*	629	*60 h*	*$\|S_{11}\|$*	*80 grid units*
[27]	*IBPSO*	1524	*24.72 h*	*$\|S_{11}\|$*	*80 grid units*
[28]	*FOA+XGBoost*	*Not given*	*13.18 h*	*$\|S_{11}\|$*	*7 parameters*
[29]	*GA+ANN*	*Not given*	*5 h*	*$\|S_{21}\|$*	*8 parameters*
[30]	*RL+DNN*	*Not given*	*40.4 h*	*$\|S_{11}\|$ and gain*	*5 parameters*
This work	*RL+E-MLPs*	704	*35.2 h*	*$\|S_{11}\|$ and gain*	*126 grid units*

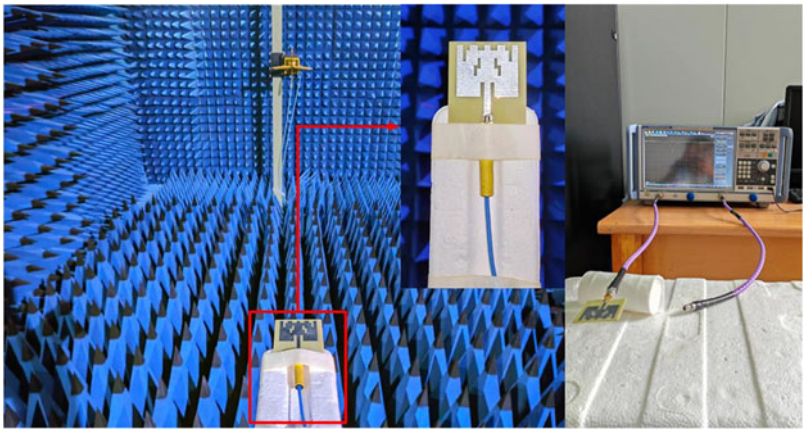


Figure 8. Measurement environment of optimized antenna.

The gain at 2.45 GHz is improved from the initial 2.55 dBi to 2.90 dBi, representing an approximately 13.7% increase. A balance between antenna impedance bandwidth and gain is achieved in the final optimization results. In addition, Fig. 7e further shows the antenna topology optimization result of each stage. As the optimization continues, the topology grid units are gradually concentrated. The final concentrated antenna topology indicates that the gain at 2.45 GHz can be ensured by the large radiating patch area and the improvement of impedance bandwidth can be achieved by the appropriate slotting. The proposed method enables a balance between those contradictory design requirements. Finally, the specific performance comparison of the optimization result in different stages is shown in Table 1.

Comparison and measurement

Comparison of the proposed method with other methods

In this work, the number of iteration is set to 100, but the optimization standard is achieved after 82 iterations. By using the optimization strategy of skipping the zero fitting value, 116 redundant EM simulation steps are effectively skipped, and 704 EM simulations are actually performed, with a total time of about 35.2 hours. The comparison of optimization results in this paper with other recent related works are shown in Table 2.

Antenna performance optimization is achieved in references [25–27] by changing the antenna predefined structural templates. Such method offers a high degree of freedom but requires thousands of EM simulations when dealing with the antenna topology problems. In contrast, references [28–30] realize the optimization through fine-tuning antenna parameters. This method requires less time to complete the antenna optimization, but its optimization efficiency is severely restricted by the initial antenna structure and the dimension of the design parameters.

Therefore, this work innovatively applies the RL to the field of antenna topology optimization. Compared with topology optimization method, this study shows an advantage in optimization efficiency. Compared with parameter optimization method, it breaks the inherent constraints of the initial topology and shows a higher degree of freedom. In addition, compared with our previous work [31, 32], the proposed method achieves the antenna multi-objective optimization with higher dimension topology data but in less optimization time.

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

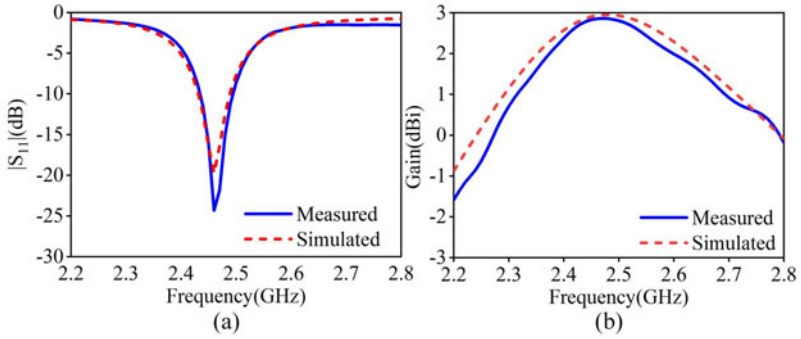


Figure 9. Measurement and simulation results of (a) $|S_{11}|$ and (b) gain.

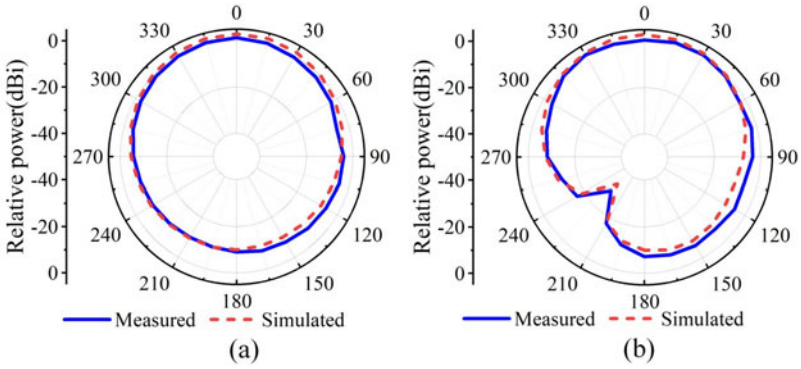


Figure 10. Measurement and simulation results of radiation pattern in (a) YOZ plane and (b) XOZ plane.

The measurement results of antenna

The experimental sample of the antenna is measured, and the measurement environment is shown in Fig. 8. In addition, Fig. 9 shows the antenna performance comparison of $|S_{11}|$ and gain between simulated and measured results. Although there are some differences which may be caused by manufacturing errors and installation differences, the antenna performance is generally consistent. Fig. 10 further shows the measured and simulated results of antenna radiation pattern in the XOZ plane and YOZ plane. The results show that the radiation patterns are also generally consistent.

Conclusion

This paper proposes an automated topology optimization method for antenna based on the E-MLPs and RL. RL is used to optimize the antenna, and E-MLPs is used to replace EM simulations. The effectiveness of this method is verified by optimizing a microstrip patch antenna. The measurement results show that the optimized antenna has an excellent performance of impedance bandwidth and gain, which meets the need of 5G communication. Additionally, compared with other antenna intelligent optimization methods, the proposed method shows a significant advantage in reducing the number of EM simulation and improving the optimization efficiency.

Acknowledgements. This paper is supported by the Yunnan Fundamental Research project [202401AT070351, 202401AT070376, 202301AV070003], Yunnan Key Laboratory of Computer Technologies Application Open Fund, the Natural Science Foundation of China [61971208], Yunnan Young Top Talents of Ten Thousands Plan [Shen Tao, Yunren Social Development No. 2018 73], College personnel training project [202101BE070001-008], and partially by the Major Science and Technology Projects in Yunnan Province under Grant [202302AG050009 and 202202AD080013].

Competing interests. The authors declare that they have no conflict of interest or competing interests in relation to this work.

References

1. **Shi D, Lian C, Cui K, Chen Y and Liu X** (2022) An intelligent antenna synthesis method based on machine learning. *IEEE Transactions on Antennas and Propagation* **70**(7), 4965–4976.
2. **Su Y** (2024) Time-domain scattering parameters-based neural network inverse model for antenna designs. *IEEE Antennas and Wireless Propagation Letters* **23**(7), 1976–1980.
3. **Wu Q, Chen W, Yu C, Wang H and Hong W** (2024) Machine-learning-assisted optimization for antenna geometry design. *IEEE Transactions on Antennas and Propagation* **72**(3), 2083–2095.
4. **Jiang Y, Yuan SSA and Sha WEI** (2024) A novel mutual coupling ANN model for MIMO antennas with physical preprocessing. *IEEE Antennas and Wireless Propagation Letters* **23**(12), 4523–4527.
5. **Minasian AA and Bird TS** (2013) Particle swarm optimization of microstrip antennas for wireless communication systems. *IEEE Transactions on Antennas and Propagation* **61**(12), 6214–6217.
6. **Hussein AH, HH Abdullah, AM Salem, S Khamis and M Nasr** (2011). Optimum design of linear antenna arrays using a hybrid MOM/GA algorithm *IEEE Antennas and Wireless Propagation Letters*. pp. 1232–1235. Vol. **10**.
7. **Li X and Luk KM** (2020) The grey wolf optimizer and its applications in electromagnetics. *IEEE Transactions on Antennas and Propagation* **68**(3), 2186–2197.
8. **He Y, Huang J, Li W, Zhang L, Wong S-W and Chen ZN** (2024) Hybrid method of artificial neural network and simulated annealing algorithm for optimizing wideband patch antennas. *IEEE Transactions on Antennas and Propagation* **72**(1), 944–949.

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press

9. **Peng F and Chen X** (2024) An efficient optimization method for HCC antennas using quantum genes based on sparse artificial neural networks. *IEEE Transactions on Antennas and Propagation* **72**(5), 3987–3999.
10. **Wang L-L, Yang X-S and Ma C-J** (2023) An efficient gradient-based hybrid parameter-topology optimization for antenna design. *IEEE Transactions on Antennas and Propagation* **71**(12), 9477–9486.
11. **Zhu E** (2024) Inverse design of multifunctional frequency-selective surface under inhomogeneous background. *IEEE Transactions on Antennas and Propagation* **72**(5), 4477–4486.
12. **Lu Y, Liu J, Zong Z and Wei Z** (2024) Multiphysics inverse design of frequency-selective surface by data-physics-driven deep neural network. *IEEE Transactions on Antennas and Propagation* **72**(11), 8739–8749.
13. **Wu D-L, Hu XJ, Chen JH, Ye LH and Li J-F** (2024) Deep learning optimization for compact dual-polarized high-isolation antenna using backpropagation algorithm. *IEEE Antennas and Wireless Propagation Letters* **23**(2), 898–902.
14. **Su Y, Yin Y, Li S, Zhao H and Yin X** (2024) Bandwidth improvement for patch antenna via knowledge-based deep reinforcement learning. *IEEE Antennas and Wireless Propagation Letters* **23**(12), 4094–4098.
15. **Yang H *et al.*** (2020). Deep reinforcement learning based intelligent reflecting surface for secure wireless communications *GLOBECOM 2020 - 2020 IEEE Global Communications Conference* **Taipei, Taiwan**. pp. 1–6.
16. **Zhou J** (2021) A trust-region parallel Bayesian optimization method for simulation-driven antenna design. *IEEE Transactions on Antennas and Propagation* **69**(7), 3966–3981.
17. **Koziel S and Pietrenko-Dabrowska A** (2024) Efficient simulation-based global antenna optimization using characteristic point method and nature-inspired metaheuristics. *IEEE Transactions on Antennas and Propagation* **72**(4), 3706–3717.
18. **Koziel S, Pietrenko-Dabrowska A and Pankiewicz B** (2024) On accelerated metaheuristic-based electromagnetic-driven design optimization of antenna structures using response features. *Electronics* **13**(2), 383.
19. **Tolstikhin I, Houlsby N, Kolesnikov A, Beyer L, Zhai X and Unterthiner T** (2021) MLP-mixer: An all-MLP architecture for vision. *Neural Information Processing Systems* **34**, 24261–24272.
20. **Zhou ZH** (2021) *Machine Learning*. Singapore: Springer Nature.
21. **Sharifani K and Amini M** (2023) Machine learning and deep learning: a review of methods and applications. *World Information Technology and Engineering Journal* **10**(7), 3897–3904.
22. **Silver D, Lever G, Heess N, Degris T, Wierstra D and Riedmiller M** (2014) Deterministic policy gradient algorithms. *PMLR* **32**(1), 387–395.
23. **Papageorgiou E, Buzo A, Pelz G and Noulis T** (2025) Deep reinforcement learning and Bayesian optimization based OpAmp design across the CMOS process space. *AEU - International Journal of Electronics and Communications* **192**, 155697.
24. **Zhang B** (2022) Ultra-wide-scanning conformal heterogeneous phased array antenna based on deep deterministic policy gradient algorithm. *IEEE Transactions on Antennas and Propagation* **70**(7), 5066–5077.
25. **Guler S, Karaosmanoglu B and Ergul O** (2016) Design, simulation, and fabrication of a novel type of inkjet-printed pixel antennas. *Progress in Electromagnetics Research Letters* **64**, 51–55.
26. **Jia X and Lu G** (2019) A hybrid Taguchi binary particle swarm optimization for antenna designs. *IEEE Antennas and Wireless Propagation Letters* **18**(8), 1581–1585.
27. **Dong J, Li Q and Deng L** (2018) Design of fragment-type antenna structure using an improved BPSO. *IEEE Transactions on Antennas and Propagation* **66**(2), 564–571.
28. **Li WT, Tang HS, Cui C, Hei YQ and Shi XW** (2020) Efficient online data-driven enhanced-xgboost method for antenna optimization. *IEEE Transactions on Antennas and Propagation* **70**(7), 4953–4964.
29. **Liu Y, Chen P, Tian JT, Xiao J and Noghanian S** (2024) Hybrid ANN-GA optimization method for minimizing the coupling in MIMO antennas. *AEU - International Journal of Electronics and Communications* **175**, 155068.
30. **Su Y, Yin Y, Li S, Zhao H and Yin X** (2024) Bandwidth improvement for patch antenna via knowledge-based deep reinforcement learning. *IEEE Antennas and Wireless Propagation Letters* **23**(12), 4094–4098.
31. **Dou J, Gong H, Chen H, Chen Y, Shen T and Song J** (2025) A reinforcement learning-based image mapping-topology optimization for antenna design. *IEEE Antennas and Wireless Propagation Letters* **24**(7), 1759–17563.
32. **Dou J, Gong H, Wei S, Chen H, Chen Y and Shen T** (2025) Bandwidth improvement for patch antenna via knowledge-based deep reinforcement learning. *ACES Journal* **40**(1), 35–41.



Jiangling Dou was born in 1985 in Jiangsu Province, China. She received her Ph.D. degree in Electromagnetic Fields and Microwave Technology from Southeast University in 2018. She is currently working at Kunming University of Science and Technology. Her research interests include electromagnetic field theory and applications.



Chenyang Zhao was born in 1999 in Henan Province, China. He is currently pursuing a master's degree at Kunming University of Science and Technology, focusing on the method of AI-assisted antenna optimization.



Jian Song (Member, IEEE) obtained his Bachelor of Engineering degree in Electronics Information Engineering from Jiangxi University of Science and Technology, Ganzhou, China. He later earned his Ph.D. degree in Electromagnetic Fields and Microwave Technology from the University of Electronic Science and Technology of China, Chengdu, China, in 2015. Between 2015 and 2019, he served as an Algorithm Engineer at Huawei Technology Co., Ltd., Shenzhen, China. In 2019, he became a Faculty Member at Kunming University of Science and Technology. His research focuses on microwave engineering and the processing of biomedical images.



Tao Shen (Senior Member, IEEE) earned his Ph.D. degree from the Illinois Institute of Technology, Chicago, IL, USA, in 2013. Presently, he holds the position of President at Yunnan Vocational and Technical College of Mechanical and Electrical Engineering. Dr. Shen has contributed to more than 20 publications in prestigious SCIE-indexed journals and leading international conferences within his research domains. His areas of expertise include image processing, artificial intelligence, and the Internet of Energy.

https://doi.org/10.1017/S1759078725102559 Published online by Cambridge University Press