Microstrip antenna modelling based on image-based convolutional neural network

Hao Fu,[1] Yubo Tian,[2,✉] Fei Meng,[2] Qing Li,[1] and Xuefeng Ren[1]

[1]*Ocean College, Jiangsu University of Science and Technology, Zhenjiang city, China*

[2]*School of Information and Communication Engineering, Guangzhou Maritime University, Guangzhou city, China*

✉Email: tianyubo@just.edu.cn

Convolutional neural networks (CNN) have a strong feature extraction ability for images and present a high level of efficiency and accuracy in object detection and image recognition. When CNN is used to model microwave devices, the existing literature generally uses its size parameters as one-dimensional (1-D) input, which does not give full play to the image-processing ability of CNN. In order to make full use of the characteristics of CNN, this letter converts the 1-D input of microwave devices into the form of an image model, that is, the 1-D input is transformed into a two-dimensional (2-D) matrix composed of 0 and 1 as the input. The image model is combined with CNN, called image-based CNN (ICNN), which establishes a deep learning surrogate model between the physical parameters and electrical properties of microwave devices and improves the accuracy and generalization ability of the model. Taking the resonant frequency of the microstrip antenna as a simulation example, modelling was carried out by the proposed ICNN and compared with the mainstream machine learning methods. The results show that the proposed method has high convergence and fitting accuracy.

Introduction: Full-wave electromagnetic simulation software has become an important tool for antenna design, however, its high-fidelity simulation is very expensive. Therefore, the use of electromagnetic simulation solvers for antenna optimization design may be limited due to their high computational cost. To solve this problem, fast and accurate machine learning (ML) surrogate models can be utilized. ML can learn features from input data and give prediction results. Up to now, many algorithms have been applied to modelling microwave devices, such as Gaussian process (GP), support vector machines (SVM), deep learning (DL) etc.

J. P. Jacobs [1] proposed an accurate modelling method for the resonant frequency of dual-band microstrip antennas based on GP regression. Jing Gao [2] proposed a semi-supervised GP co-training method based on different kernel functions for antenna resonant frequency modelling, and a small number of labelled samples and unlabelled samples were combined to continuously improve the accuracy of the model. Xie Zheng [3] has proposed a monopole antenna design based on the progressive GP.

Jing Gao [2] proposed a semi-supervised co-training algorithm based on GP and SVM. Dan Shi [4] proposed a smart antenna synthesis method, which automatically selects appropriate antenna types according to antenna performance requirements and uses SVMs to provide optimal geometric parameters.

Jing Jin [5] introduced deep neural network (DNN) method into the field of high-dimensional microwave modelling and proposed a DNN technology for high-dimensional microwave modelling and its application in microwave filter parameter extraction. Andrea Massa [6] mentioned that solving the forward electromagnetic (EM) problem in real-time is still very challenging due to the complexity of the antenna numerical problem, and DNNs are rapidly emerging as promising candidates to significantly speed up standard EM solvers.

Due to the addition of local receptive fields, parameter sharing and sparse weights, convolutional neural network (CNN) has the characteristics of translation and scale invariance compared with other neural networks, which makes CNN can effectively extract features and is more suitable for image data processing [7]. Zhang [8] proposed an algorithm combining particleswarm optimization (PSO) and CNN for the design of high-dimensional and compact pixel microstrip antennas. Compared with the results of traditional simulation software, the optimization results were quickly obtained. Jacobs [9] proposed an accurate modelling method for the resonant frequency of a dual-band pixel microstrip antenna based on CNN regression, which takes the entire pixelized surface of the antenna as input. Hai-Ying [10] proposed a method for parametric modelling of ultra-wideband (UWB) antennas using CNN. Given an image of the antenna, CNN can quickly predict its electromagnetic response and speed up the design process. The model input proposed by Zhang is a 1-D input of the antenna size parameters and the model input proposed by Jacobs and Hai-Ying are both images of the antenna itself.

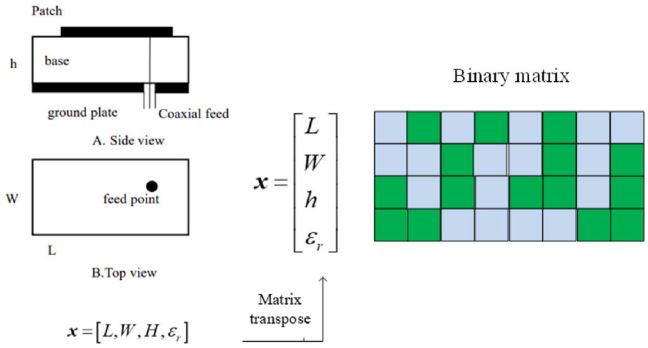

Fig. 1 *Construction of image model of the rectangular microstrip antenna.*

The models established in the existing literature generally take the antenna size parameters as one-dimensional (1-D) input, which does not give full play to the image processing ability of CNN. Therefore, in order to improve the modelling accuracy of CNN, the microstrip antenna size parameters are modelled as an image model, that is, the 1-D model input is mapped to a two-dimensional (2-D) matrix composed of 1 and 0, similar to black and white image data, to form an image model based CNN, and its prediction performance is compared with the results obtained by traditional ML methods in this letter.

Model construction: In order to describe the image model defined in this letter, a rectangular microstrip antenna (RMSA) is used for illustration in detail. The length of the antenna is L, the width is W, the thickness of the dielectric layer is h and the relative dielectric constant is ε_r.

When CNN models the resonant frequency of the RMSA, its input parameters can be expressed as $x = [L, W, h, \varepsilon_r]$. Considering that 1-D CNN can not give full play to the characteristics of the CNN model, the form of a 2-D image model is constructed, as shown in Figure 1.

To be specific, we transpose the input $\boldsymbol{x}$ of the antenna to a binary string of either 0 or 1, and its value is denoted as

$$\mathbf{x_i} = (x_i)_{min} + \frac{\sum_{k=0}^{m-1} a_k 2^k}{2^m - 1}\left((x_i)_{max} - (x_i)_{min}\right) \tag{1}$$

In the formula, $(x_i)_{min}$ represents the minimum value within the range of x_i, $(x_i)_{max}$ represents the maximum value within the range of x_i, m is the length of the binary string, and a_k is 1 or 0, which represents the existence of this position.

In the above equation, if $\sum_{k=0}^{m-1} a_k 2^k$ is regarded as A, it can be concluded that:

$$A = \frac{x_i - (x_i)_{min}}{(x_i)_{max} - (x_i)_{min}}\left(2^m - 1\right) \tag{2}$$

where, the meaning of x_i, $(x_i)_{min}$, $(x_i)_{max}$and m is the same as that of Equation (1). Thus, the decimal value represented by A can be obtained, and the corresponding relationship between the decimal value representing the antenna structure parameters (i.e. the problem space) and the binary string represented by Figure 1 (i.e. the image model space) can be formed. The image model constructed in the form of 2-D binary string after conversion is used as the input of CNN to form Image-based CNN (ICNN), which can give full play to the visual processing characteristics of CNN, and make the model more accurate compared with one-dimensional CNN.

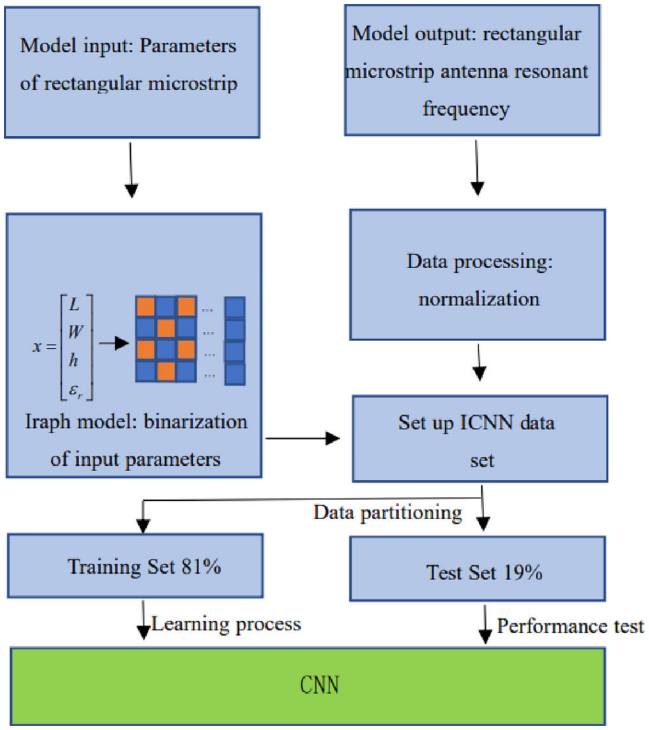

Fig. 2 *Flowchart of the proposed image-based convolutional neural network.*

In the process of transformation between the problem space and the image model space, the following three specifications should be satisfied:

1. Completeness: Every 1-D vector in the problem space can be represented by a 2-D binary image in the image model space;
2. Soundness: The 2-D binary images in the image model space correspond to all 1-D vectors in the problem space;
3. Non-redundancy: It means one-to-one relationship between the 1-D vector in the problem space and the 2-D binary image in the image model space.

Simulation experiment: In this part, the resonant frequency of the RMSA is modelled by the proposed ICNN.

Step 1 Traditionally, when the CNN is used for modelling the resonant frequency of the RMSA, the input is a 1D vector $x = [L, W, h, \varepsilon_r]$ composed of four elements. In the proposed ICNN, if the length of the binary string of each element is controlled to be 10, then a two-dimensional matrix of 4×10 can be generated after the processing of the image model. Totally, there are 33 groups of data, which are divided into two parts, of which 27 groups (81%) are used as training samples, and six groups (19%) are used as testing samples. Figure 2 is the flowchart of the ICNN used for modelling the resonant frequency of the RMSA.

Step 2 Determining the hyperparameters of the ICNN is a complex engineering problem that has a significant impact on the accuracy and generalization of the trained neural network, which is briefly explained below.

1. Activation functions: To enhance the ability to solve complex problems, activation functions in neural networks are introduced for nonlinear transformations. For the ICNN, we use the Relu function.
2. Loss function: The loss function ensures the best performance of the ICNN on the training dataset. In this experiment, the root mean square error (RMSE) and the average percentage error (APE) are used to measure the deviation between the regression output value and the realistic target value, and the coefficient of determination R^2 is used to measure the degree of fit of the regression model. Their formulas are as follows.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - Y_i)^2} \tag{3}$$

Table 1. Model hyperparameters.

Parameters	Value
Optimizer chooses	Sgd/Adam/Rmsprop
Convolution kernels	3×5 and 2×2
Number of neurons	20
Batch size	3
Dropout	0.01/0.05/0.1/0.2/0.3/0.4
Epochs	200

Table 2. The structure of the convolutional neural networks.

	Input size	Output size	Weight matrix size
Convolutional layer 1	$4 \times 10 \times 1$	$4 \times 8 \times 2$	$2 \times 5 \times 2$
Convolutional layer 2	$4 \times 8 \times 2$	$4 \times 6 \times 4$	$2 \times 5 \times 4$
Convolutional layer 3	$4 \times 6 \times 4$	$3 \times 5 \times 8$	$2 \times 2 \times 8$
Fully connected layer 1	120	100	100×120
Fully connected layer 2	100	20	20×100
Output layer	20	1	1×20

$$APE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - Y_i|}{|y_i|} * 100 \tag{4}$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{5}$$

Where $SS_{res} = \sum_{i=1}^{N}(y_i - Y_i)^2$, $SS_{tot} = \sum_{i=1}^{N}(y_i - \overline{y_i})^2$, $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$, y_i is the real value, Y_i is the predicted value, and N is the number of samples.

3. Feature optimizer: The feature optimizer is used to find the optimal value for the loss function. Adaptive moment estimation (ADAM) calculates the first moment estimate and the second moment estimate of the gradient to design independent adaptive learning rates for different parameters. Stochastic gradient descent (SGD) is a gradient-based optimization algorithm that can help find the parameters that minimize the loss function. Root mean square propagation (RMSProp) uses a hyperparameter to smooth the momentum term and the squared gradient. It makes greater progress in the direction of the optimal sulution of the objective function, which can reduce the problem of too large gradient swing during the iteration.
4. Number of epochs and batch size: An epoch is the one when the entire data set is passed forward and backward through the neural network once. If the entire data set cannot be passed into the algorithm at once, it must be split into several mini-batches. Batches size refers to the total number of training samples presnet in a single batch.

According to the specific situation of the experiment and some prior knowledge, the model hyperparameters are determined in Table 1. The CNN is composed of three convolutional layers and two fully connected layers. On the basis of setting the hyperparameters, the structure of the convolutional neural network in the figure model is shown in Table 2.

Step 3 The ICNN model is trained according to the existing model hyperparameters using the training samples. To ensure the fairness of the experiment and the reliability of the results, each experiment was independently run 20 times to obtain the optimal value. The resonant frequencies of the results of the RMSA measured by experiments under different drops and optimizers are shown in Tables 3–5.

Comparison and analysis: In this section, the model performance is verified by comparing the experimental results.

Case 1: Comparison of results of resonant frequencies of RMSA measured by different drop probability and optimizers.

The parameter settings of the proposed ICNN are shown in Table 1, and the experimental results are shown in Tables 3–5. It can be seen that different dropout probabilities and optimizers have significant effects on the experimental results. Within the dropout probability range, when the

Table 3. APE of resonant frequency of rectangular microstrip antenna measured by different dropout probability and optimizers.

	Optimizer		
Dropout probability	Sgd	Adam	Rmsprop
0.01	1.2109	0.63047	1.3618
0.05	0.7728	09865	1.242
0.1	0.48210	1.2851	1.3187
0.2	0.6820	1.2163	1.3297
0.3	1.4585	1.7366	2.1876
0.4	2.8103	2.1548	1.5058

Table 4. RMSE of resonant frequency of rectangular microstrip antenna measured by different dropout probability and optimizers.

	Optimizer		
Dropout probability	Sgd	Adam	Rmsprop
0.01	107.16	57.90	109.69
0.05	57.57	86.22	112.60
0.1	54.43	121.59	97.15
0.2	74.89	85.76	100.06
0.3	121.91	135.76	172.40
0.4	220.05	182.52	139.94

Table 5. R^2 *of resonant frequency of rectangular microstrip antenna measured by different dropout probability and optimizers.*

	Optimizer		
Dropout probability	Sgd	Adam	Rmsprop
0.01	0.9721	0.9732	0.9594
0.05	0.9809	0.9670	0.9306
0.1	0.9977	0.9510	0.9262
0.2	0.9661	0.9253	0.8983
0.3	0.9108	0.9139	0.8690
0.4	0.8402	0.8464	0.8654

Table 6. Predicted results of different modelling methods for the rectangular microstrip antenna.

f (MHz)	1D CNN	GP [11]	DKL52 [11]	ICNN
3200	3216.23	3196.581	3196.579	3149.43
3580	3597.51	3582.141	3584.779	3580.81
4805	4826.13	4830.272	4823.193	4789.17
5100	5122.44	5145.394	5088.715	5092.45
6200	6205.49	6198.311	6197.412	6162.057
8450	8142.09	8251.843	8239.441	8504.42
APE	0.9348	0.6400	0.519	**0.4821**

dropout probability is equal to 0.1, the model convergence performance reaches the optimum, and then the convergence performance gradually decreases with the increase of dropout probability. All three optimizers have certain convergence performance, but SGD has higher convergence accuracy than the others.

Case 2: Comparison of the results of resonant frequencies of RMSA measured by different models.

In this case, several models are listed in Table 6. It shows the comparison results of the ICNN, GP, deep kernel learning (DKL52) [11] and 1-D CNN for modelling and predicting the resonant frequency of the RMSA. It can be seen that the accuracy of the ICNN is increased by 48% compared with 1D CNN, 24% compared with APE compared with GP, and 18% compared with DKL52. The results show that the ICNN has higher modelling accuracy and better generalization ability. According to Table 6, the hyperparameters of 1-D CNN model in Table 6 are shown in Table 7. The real value of RMSA resonance frequency and the predicted value of the ICNN are shown in Figure 3.

Table 7. 1D convolutional neural networks model hyperparameters.

Parameters	Value
Optimizer chooses	sgd
Convolution kernels	2×1
Batch size	3
Dropout	0.1
Epochs	200

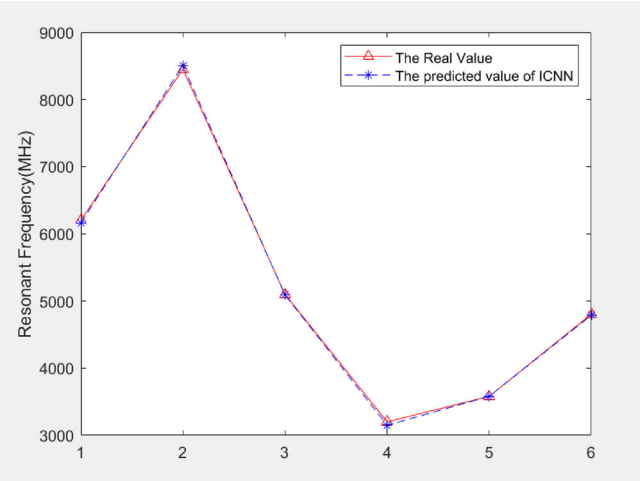

Fig. 3 *The real value of rectangular microstrip antenna resonance frequency and the predicted value of image-based convolutional neural networks.*

Conclusion: In this paper, we propose a new data processing technique-image model, which transforms the original one-dimensional data into two-dimensional images, to enhance the powerful feature extraction ability of CNN modelling. Through the resonant frequency modelling example of a microstrip antenna, the performance of the algorithm is tested. The experimental results show that the performance of the proposed surrogate model is superior to other modelling methods. The method presented in this letter can be easily extended to other microwave engineering.

Author contributions: **Hao Fu**: Conceptualization; Data curation; investigation; methodology; validation; visualization; writing original draft; and writing review and editing. **Yubo Tian**: Data curation; investigation; methodology; investigation; formal analysis; funding acquisition; project administration; validation. **Fei Meng**: Data curation; formal analysis; supervision; investigation; resources; visualization. **Qing Li**: Methodology; investigation; and resources. **Xuefeng Ren**: Methodology; investigation; and resources.

Acknowledgements: This work was supported by the Natural Science Foundation of Guangdong Province of China under Grant No. 2023A1515011272, the Special Project in Key Fields of Guangdong Universities of China under No. 2022ZDZX1020, and the Tertiary Education Scientific research project of Guangzhou Municipal Education Bureau of China under No. 202234598.

Conflict of interest statement: The authors declare no conflicts of interest.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: *17 May 2023* Accepted: *31 July 2023*
doi: 10.1049/ell2.12910

References

1 Jacobs, J.P.: Efficient resonant frequency modeling for dual-band microstrip antennas by Gaussian process regression. *IEEE Antennas Wirel. Propag. Lett.* **14**, 337–341 (2014)
2 Gao, J., Tian, Y., Chen, X.: Antenna optimization based on co-training algorithm of Gaussian process and support vector machine. *IEEE Access* **8**, 211380–211390 (2020). https://doi.org/10.1109/ACCESS.2020.3039269
3 Zheng, X., Meng, F., Tian, Y.,Zhang, X.: Design of monopole antennasbased on progressive Gaussian process. Int. J. Microwave Wireless Technolog. **15**(2), 1–8 (2022). https://doi.org/10.1017/S1759078722000125
4 Shi, D., Lian, C., Cui, K., Chen, Y., Liu, X.: An intelligent antenna synthesis method based on machine learning. *IEEE Trans. Antennas Propag.* **70**(7), 4965–4976 (2022). https://doi.org/10.1109/TAP.2022.3182693
5 Jin, J., Zhang, C., Feng, F., Na, W., Ma, J., Zhang, Q.-J.: Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters. *IEEE Trans. Microwave Theory Tech.* **67**(10), 4140–4155 (2019). https://doi.org/10.1109/TMTT.2019.2932738
6 Massa, A., Marcantonio, D., Chen, X., Li, M., Salucci, M.: DNNs as applied to electromagnetics, antennas, and propagation–A review. *IEEE Antennas Wirel. Propag. Lett.* **18**(11), 2225–2229 (2019) https://doi.org/10.1109/LAWP.2019.2916369
7 McCann, M.T., Jin, K.H., Unser, M.: Convolutional neural networks for inverse problems in imaging: A review. *IEEE Signal Process Mag.* **34**(6), 85–95 (2017)
8 Zhang, X., Tian, Y., Zheng, X.: Optimal design of fragment-type antenna structure based on PSO-CNN[C], In: 2019 International Applied Computational Electromagnetics Society Symposium-China (ACES), pp. 1–2. IEEE, Piscataway, NJ (2019)
9 Jacobs, J.P.: Accurate modeling by convolutional neural-network regression of resonant frequencies of dual-band pixelated microstrip antenna. *IEEE Antennas Wirel. Propag. Lett.* **20**(12), 2417–2421 (2021)
10 Luo, H.-Y., Hong, Y., Lv, Y.-H., Shao, W.: Parametric modeling of UWB antennas using convolutional neural networks, In: 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting, pp. 2055–2056. IEEE, Piscataway, NJ (2020). https://doi.org/10.1109/IEEECONF35879.2020.9329697
11 Han, S., Tian, Y., Ding, W., Li, P.: Resonant frequency modeling of microstrip antenna based on deep kernel learning. *IEEE Access* **9**, 39067–39076 (2021) https://doi.org/10.1109/ACCESS.2021.3062940