Journal of Applied Physics | ARTICLE | scitation.org/journal/jap

An electromagnetic parameter retrieval method based on deep learning

Cite as: J. Appl. Phys. **127**, 224902 (2020); doi: 10.1063/5.0005455
Submitted: 24 February 2020 · Accepted: 25 May 2020 ·
Published Online: 8 June 2020

View Online Export Citation CrossMark

Siqi Huang,[1] Zilong Cao,[2] Helin Yang,[1,a)] Zhaoyang Shen,[1] and Xiaoxia Ding[1]

AFFILIATIONS

[1]School of Physical Science and Technology, Central China Normal University (CCNU), Wuhan 430079, China
[2]School of Computer Science and Technology, Huazhong University of Science and Technology (HUST), Wuhan 430074, China

[a)]**Author to whom correspondence should be addressed:** emyang@mail.ccnu.edu.cn

ABSTRACT

The electromagnetic parameter retrieval method based on scattering (*S*) parameters is widely used because *S* parameters are easy to obtain. As a classic *S* parameters retrieval method, the Nicolson–Ross–Weir (NRW) algorithm has problems such as half-wave resonance, phase angle jump, and multivaluedness. In this work, an electromagnetic parameter retrieval method based on deep learning is proposed, which aims to solve the multivaluedness problem in the traditional NRW method. The method we proposed is suitable for retrieval of inhomogeneous medium that satisfies the theory of effective medium. The simulation shows that the electromagnetic parameter retrieval method based on deep learning has high calculation accuracy without the multivaluedness problem.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0005455

I. INTRODUCTION

It is very convenient to replace the composite structural material by a homogeneous medium.[1] This equivalent can be easily achieved when the applied fields are static or have spatial variation on a scale significantly larger than the scale of the local inhomogeneity, in which case the composite is said to form an effective medium. Ideally, there is no difference in the electromagnetic (EM) response observed between the effective medium and the composite material.

There are many methods to obtain the electromagnetic parameter of an effective medium, such as the field average method,[2] curve fitting method,[3] dispersion equation method,[4] and scattering (*S*) parameters extraction method.[5–10] The *S* parameter-based retrieval method is more widely used because the *S* parameters are easier to obtain. As a *S* parameter-based retrieval method, the traditional Nicolson–Ross–Weir (NRW) method[11–13] has the characteristics of a wide frequency band and high precision and is widely used to extract the permittivity and permeability of a homogeneous sample of isotropic material under specified illumination conditions. Traditional NRW methods have problems such as half-wave resonance, phase angle transitions, and multivaluedness.[14] Among them, the primary solution to the problem of multivaluedness is the group delay method.[12] However, it requires a lot of additional data processing programs, and the calculation operation is complicated. Therefore, it is important to find a simple and automatic retrieval method.

Deep learning[15] is a branch of machine learning. It is an algorithm that tries to abstract the data at a high level using multiple processing layers that contain complex structures or consist of multiple non-linear transformations. At present, several deep learning frameworks, such as convolutional neural networks, deep belief networks, and recursive neural networks, have been applied in the fields of computer vision, speech recognition, natural language processing, audio recognition, and bioinformatics and have obtained excellent effect. In Ref. 16, the authors also used deep learning techniques to assist in the design of metamaterials to reduce the workload of operators.

In this paper, we propose an electromagnetic parameter retrieval method based on deep learning to solve the multivaluedness problem in traditional NRW methods. To our knowledge, this is the first attempt to apply deep learning to electromagnetic parameter retrieval. In the method proposed in this paper, the *S* parameters need to be reduced in dimension first and then combined with the corresponding electromagnetic parameter to form a sample dataset used to train the deep learning model. The experimental results show that the model we proposed can find the inherent law between the *S* parameters and their electromagnetic parameter. It can automatically output an electromagnetic

Published under license by AIP Publishing.

parameter according to the S parameters, and there is no multivaluedness problem in the traditional NRW method.

II. DEEP LEARNING MODEL

In this section, we first discuss the multivaluedness problem in the NRW method and then describe the process of establishing a deep learning model. Before modeling, we first get enough sample data through the Visual Basic (VBA) scripting language in Cell Signaling Technology (CST), and then use the cepstrum transform to compress the S_{11} and S_{21} parameters to 100 dimensions and prove that the features after dimensionality reduction can accurately restore the original data. In order to make full use of the sequence information of the data, we use a gated recurrent unit (GRU) neural network[17] to establish an retrieval model. Finally, we train the network model with the generated samples, and the trained model can automatically retrieve the permittivity and permeability by entering S parameters.

A. Multivaluedness problem in NRW methods

The NRW method calculates the electromagnetic parameter of the material by measuring the reflection and transmission response of the transmission line (waveguide or coaxial line) filled with the measured medium. According to the electromagnetic field theory, when a plane wave is projected from free space onto a medium surface of infinite thickness, energy reflection occurs. The reflection coefficient is

$$\Gamma = \frac{Z_S - Z_0}{Z_S + Z_0} = \frac{\sqrt{\frac{\mu_r}{\varepsilon_r}} - 1}{\sqrt{\frac{\mu_r}{\varepsilon_r}} + 1}. \tag{1}$$

Coaxial air lines are often used as samplers in actual tests, as shown in Fig. 1. Assume that the thickness of the measured material is d, and the relative permittivity and permeability are ε_r and μ_r, respectively. There are multiple reflections/transmissions on both interfaces of air and the sample. The vector network analyzer adopts the frequency sweep method to measure the measured reflection (S_{11}) and transmission (S_{21}) at the two ports of the coaxial transmission line. It is the superposition of the electromagnetic wave after multiple reflections and transmissions. The relationship between the S parameters of the two ports and the reflection coefficient Γ and propagation factor T is

$$S_{11}(\omega) = \frac{(1 - T^2)\Gamma}{1 - \Gamma^2 T^2}, \tag{2}$$

$$S_{21}(\omega) = \frac{(1 - \Gamma^2)T}{1 - \Gamma^2 T^2}. \tag{3}$$

Here, $\omega = 2\pi f$ is the working angular frequency, and T is the propagation factor given by

$$T = e^{-\gamma d}, \tag{4}$$

where $\gamma = \gamma_0\sqrt{\varepsilon_r \mu_r}$ and $\gamma_0 = j\omega\sqrt{\varepsilon_0 \mu_0}$ are the propagation constants of the sample area and air, respectively. For γ_0, ε_0 and μ_0 are the permittivity and permeability of air, respectively.

According to the NRW algorithm, let

$$V_1 = S_{21} + S_{11}, \tag{5}$$

$$V_2 = S_{21} - S_{11}. \tag{6}$$

Then,

$$X = \frac{1 - V_1 V_2}{V_1 - V_2} = \frac{1 - \left(S_{21}^2 - S_{11}^2\right)}{2S_{11}}, \tag{7}$$

$$\Gamma = X \pm \sqrt{X^2 - 1}(|T| \le 1), \tag{8}$$

$$T = \frac{V_1 - \Gamma}{1 - V_1\Gamma}. \tag{9}$$

If

$$c_1 = \left(\frac{1 + \Gamma}{1 - \Gamma}\right)^2 = \frac{\mu_r}{\varepsilon_r}, \tag{10}$$

$$c_2 = -\left(\frac{c}{\omega d}\ln\left(\frac{1}{T}\right)\right)^2 = \mu_r \varepsilon_r, \tag{11}$$

then

$$\mu_r = \sqrt{c_1 c_2}, \tag{12}$$

$$\varepsilon_r = \sqrt{\frac{c_2}{c_1}}. \tag{13}$$

Theoretically, ε_r and μ_r can be finally found by S_{11} and S_{21}. But in Eq. (11), the value of $\ln(1/T)$ is not unique, that is,

$$\ln(1/T) = \ln(\alpha + j\beta) = A + jB = A + jB + j2n\pi. \tag{14}$$

In Eq. (14), n is a natural number, so there is an infinite number of values of $\ln(1/T)$, and their imaginary parts differ by

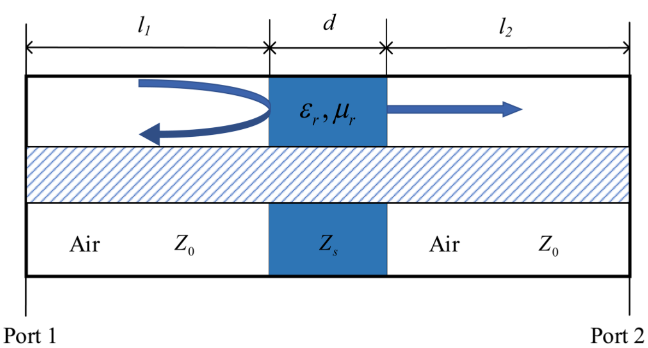

FIG. 1. Coaxial fixture used in the NRW method.

Published under license by AIP Publishing.

$2n\pi$. When performing a mathematical operation, the solution within $[-\pi, +\pi]$ is generally given directly. When the imaginary part exceeds $\pm\pi$, it jumps directly to the value within $[-\pi, +\pi]$, and the n value determined therefrom is random. This is the multivaluedness problem of the NRW method. Therefore, we try to use neural networks to derive the relationship between S parameters and electromagnetic parameters, in order to solve the problem of multivaluedness under the premise of simplicity and ease of use.

B. Data preprocessing

1. Collection of data

Training a deep neural network requires a large number of samples. In order to build a dataset containing the correspondence between S parameters, permittivity and permeability, we used CST MICROWAVE STUDIO (MWS) to generate simulation data. In order to generate simulation data in batches, we used the Visual Basic (VBA) language integrated in CST and Python scripts to automatically run the simulation programs in batches. We use FR4 (loss-free) as the dielectric substrate which has a permittivity of 4.3 and a permeability of 1 (at 10 GHz). It is worth noting that we do not limit the acquisition method of S parameters, so we used a simple S parameter scanning method instead of the coaxial line in NRW, which can reduce the difficulty of actual operation. As shown in Fig. 2, we use a 2 mm thick substrate and place two ports parallel to the dielectric plate to simulate the S parameters. Then, we modify the permittivity and permeability of the template

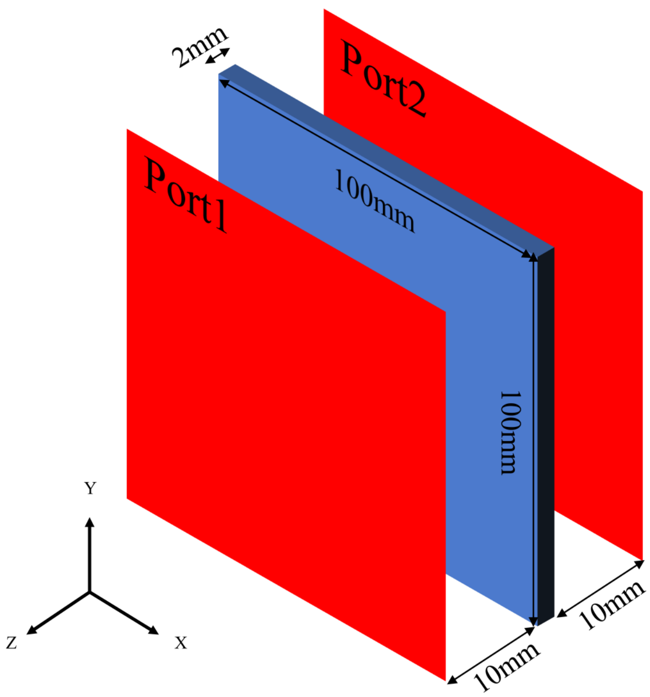

FIG. 2. Simulation project settings.

material to generate different materials. We set the value of the permittivity from −10 to 10 in steps of 0.1 (without 0) and set the value of the permeability from −5 to 5 with a step size of 0.1 (without 0). In this way, a total of 20 000 sets of data can be generated.

Finally, we stored the S parameters data as a table, including a total of 20 000 rows of data, each row of data has a total of 4002 columns, the first two columns are the permittivity and permeability, and the next 4000 columns are the S parameters at different frequencies (including S_{11}, S_{21}, S_{22}, S_{12}, each S parameters is 1000 dimensions). Because the material is symmetric, S_{11} is exactly the same as S_{22}, and S_{21} is exactly the same as S_{12}, only S_{11} and S_{21} are used in our experiments.

The detailed data collection process is shown in Fig. 3 and the steps are described as follows:

1. **Generate CST projects in batches.** First, build CST projects in batches through VBA statements in the CST and set the materials in the project to different electromagnetic parameters.
2. **Start simulation tasks in batches.** CST provides a method to execute the project through command line statements. We use a Python script to execute CST simulation tasks in batches to obtain simulation results.
3. **Read the *S* parameters data.** After the simulation is completed, we need to use VBA scripts to extract the S parameters data results in each project and write the results of each project to a text file separately.
4. **Summarized into a tabular file.** We use a Python script to read the text data containing S parameters in all projects and combine the data with the corresponding permittivity and permeability into a csv table file.

2. Feature extraction

After the sample data are obtained, if input the S parameters to the deep neural network model directly, then the model will be too complicated and the expected results will not be obtained due to the high dimensionality of the S parameters. Therefore, we need to extract the main features of the S parameters to reduce the dimensionality of the data. When measuring S parameters in real environment, because the voltage or current is unstable, EM waves caused by interference sources may generate noise during the measurement. Since these types of noise always follow the Gaussian distribution, data can be Gaussian filtered in advance to reduce noise interference.

The physical meaning of the S parameters is the power ratio, and it changes with frequency. In order to reduce the dimensionality of S parameters, we use the cepstrum method in speech processing which can be used to extract features of frequency domain data. Figure 4 shows the process of extracting Mel Frequency Cepstral Coefficients (MFCCs) features of speech signal which are the features widely used in speech processing. In the MFCC extraction process, after converting the voice signal to the frequency domain using Discrete Fourier transform (DFT), steps 6 and 7 are cepstrum analysis steps used to extract the features of the voice signal in the frequency domain. So, we can reduce the dimensionality of S parameters by using cepstrum analysis.

Published under license by AIP Publishing.

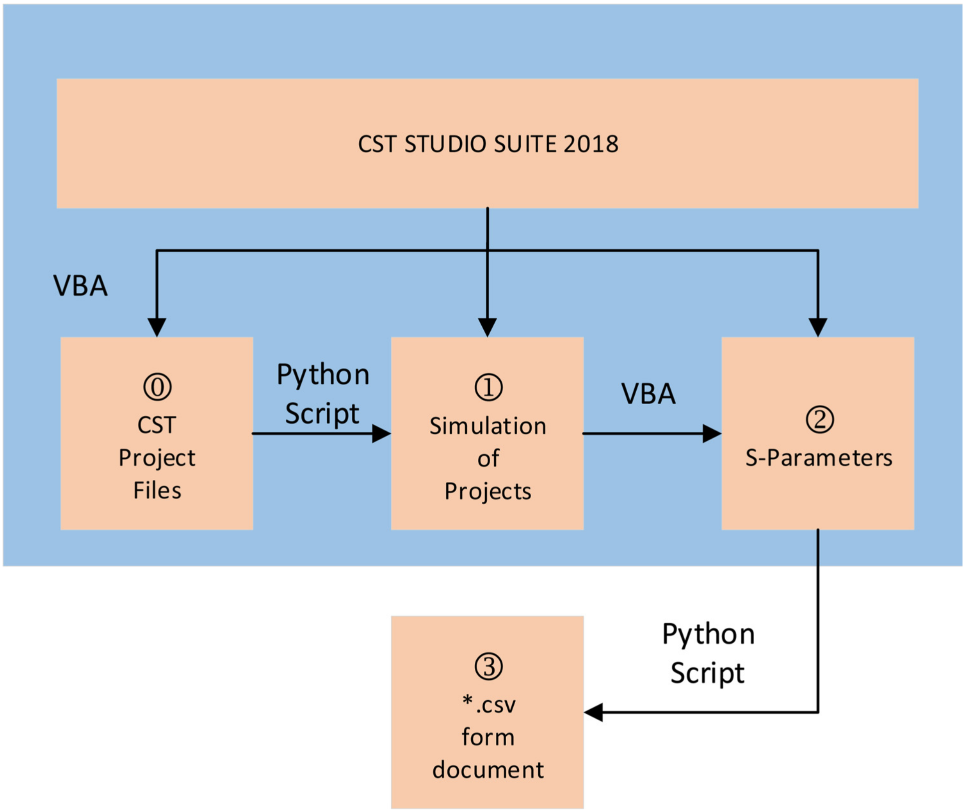

FIG. 3. Data collection process.

Cepstrum is defined as the Inverse Discrete Fourier Transform (IDFT) of the logarithm of the spectrum which represent a complicated rearrangement of time-frequency transforms. It should be noted that doing IDFT on the logarithmic domain of the frequency spectrum can be described as representing signals on the quefrency-axis which is a pseudo-time axis and expressed typically in the unit seconds. The key information in the original spectrum is the envelope of the spectrum, which is reflected in the low

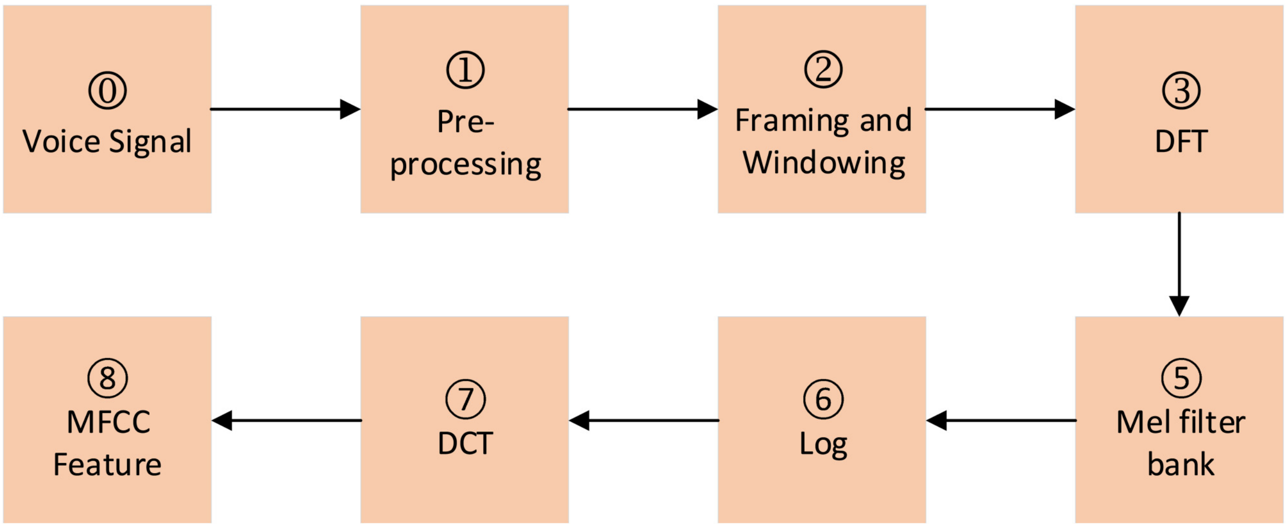

FIG. 4. MFCC feature extraction process in speech processing.

Published under license by AIP Publishing.

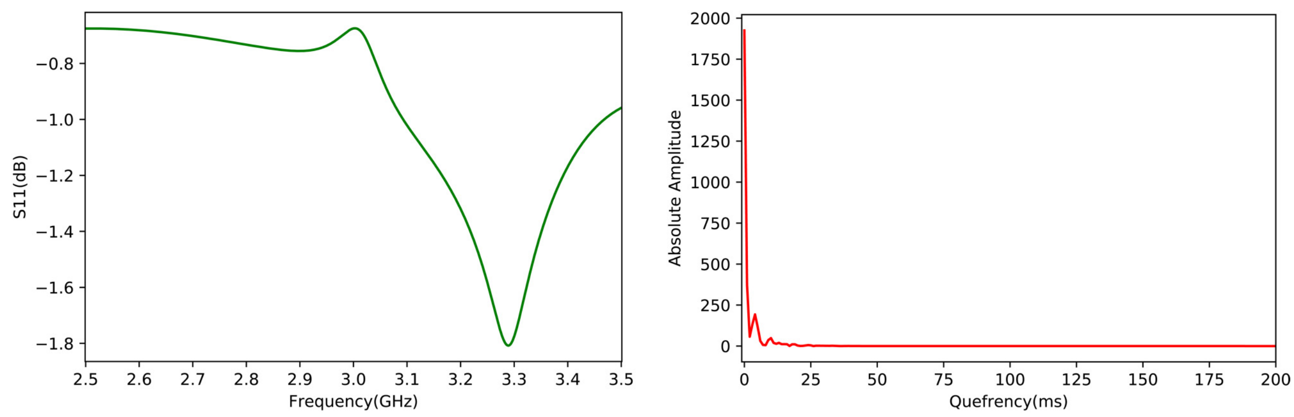

FIG. 5. DCT results of S parameters. (a) S_{11} parameters of the FR-4 dielectric board. (b) DCT results of the S_{11} parameters.

quefrencies. We can get the main information by passing through a low-pass filter. This process of cepstrum analysis can be summarized as follows:

(a) Assume the S parameters data are expressed as $X[k] = H[k]E[k]$, where $H[k]$ is the envelope of the spectrum and $E[k]$ is the details of the spectrum.
(b) Logarithm on both sides of the equation: $\log X[k] = \log H[k] + \log E[k]$.
(c) Take IDFT on both sides of the equation to get: $x[k] = h[k] + e[k]$, where $x[k]$ is the cepstrum.
(d) Finally, the low-pass filter is used to obtain $h[k]$, which describes the envelope of the spectrum.

In step (c), the result is a complex number if IDFT is used. In order to keep the results in the real number domain, IDFT is replaced with discrete cosine transform (DCT) in MFCC, which also facilitates data processing and storage. Figure 5 shows the S_{11} parameters of the FR-4 dielectric board and its DCT results. For most samples, the information is concentrated in the first 100 points, so we extract the first 100 points as our compressed features. After this, the features have been reduced from the initial 1000 dimension down to 100 dimensions and they can can restore most of the original input.

In order to prove that the compressed features contain most of the information in the original signal, we use inverse DCT (IDCT) on the compressed 100-dimensional features to restore the original signal. It can be seen from Fig. 6 that the restored signal mostly coincides with the original signal (the data are standardized).

C. The establishment of a deep neural network model

In order to make full use of the sequence information of the data, we use the GRU network model, which is an optimized recurrent neural network (RNN). The original RNN is prone to gradient attenuation or explosion problems.[18] Although gradient cropping can deal with gradient explosions, they cannot solve the problem of gradient attenuation. The proposed gated recurrent neural network is precisely captures the dependency of the longer time step distance in the time series by controlling the flow of information through the gates. Among them, gated recurrent unit (GRU) is a commonly used gated recurrent neural network. GRU introduces the concepts of reset gate and update gate. As shown in Fig. 7, the reset gate helps capture short-term dependencies in the time series, and the update gate helps capture long-term dependencies. Therefore, the GRU is better at capturing dependencies for time series with large time step distances.

We combine the 100-dimensional features of S_{11} and S_{21} into a 100*2 tensor as the input of the GRU network. After the GRU network, we added a layer of fully connected neural network. The output of the network is the permittivity and permeability. The final network model $\phi_A(\boldsymbol{\theta}_a)$ is shown in Fig. 8.

Because the value of the S parameters is the attenuation ratio, the value range is 0–1, and the value after taking log is less than 0. In order to make the network update faster, it is necessary to normalize the data with an average of 0 and a variance of 1. Therefore,

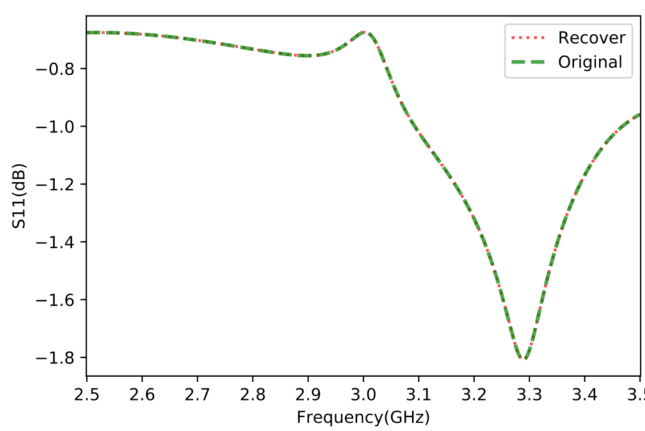

FIG. 6. S_{11} parameters restoration through extracted features.

Published under license by AIP Publishing.

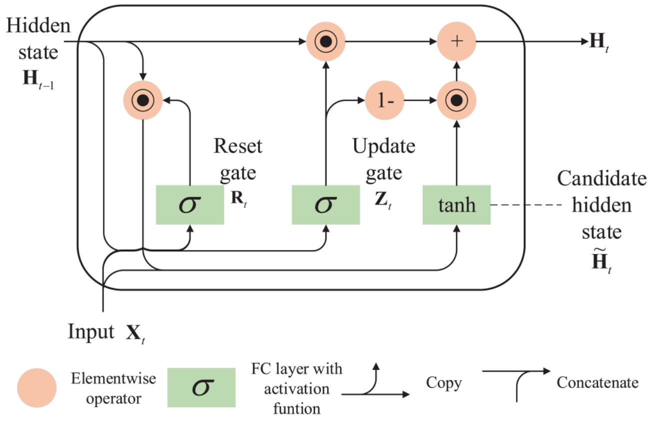

FIG. 7. Hidden state computation in a GRU.

we added batch normalization (BN)[19] to the network which reduces the offset of the hidden unit value (covariance offset). There are many benefits to using BN, including reducing the network's dependence on parameter initialization, making model training faster, achieving a regularization effect, and increasing network generalization. In order to further improve the generalization of the network, we also add an L2 regularization term to the network. Finally, we randomly split 80% of the dataset as the training dataset and the remaining 20% as the test dataset.

In order to speed up training, we use the GPU for computational acceleration. Pytorch[20] provides a GPU interface for deep learning algorithms, which allows programs to be executed in parallel in the GPU, significantly reduce the training time.

After training the deep learning model, we can measure the S parameters of the dielectric plate with unknown electromagnetic

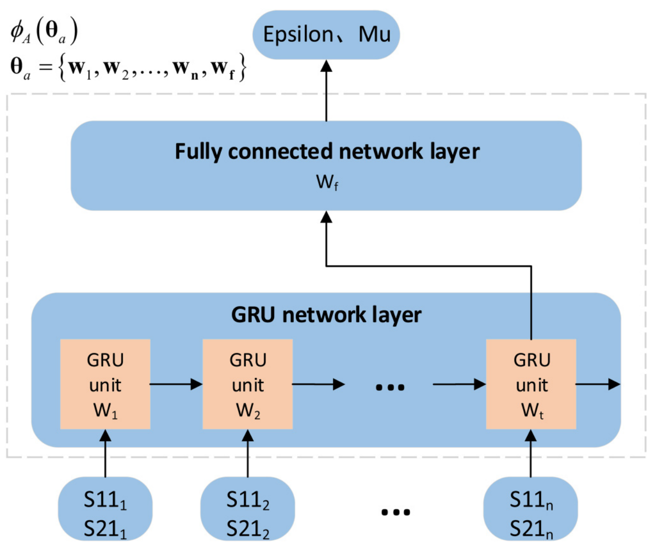

FIG. 8. GRU based inverse network $\phi_A(\boldsymbol{\theta}_a)$.

TABLE I. Hyperparameters in deep neural networks.

Hyperparameters	Values
Input size	2
Number of hidden layer	2
Number of classes	2
Hidden size	256
Batch size	200
Learning rate	0.003
Activation function	ReLu
Optimizer	Adam
Loss function	MSE

parameters by the method shown in Fig. 2 and get the permittivity and permeability by inputting the S parameters into the model.

III. RESULTS AND DISCUSSION

Our dataset is generated by CST STUDIO SUITE 2018 under Windows 7. The algorithm is deployed under Ubuntu 16.04 and the hardware configuration of the computer is CPU E5-2650 v3, GPU GTX 1080TI, and memory 128G. The deep learning algorithm is implemented on the Pytorch platform 1.0, and the Python version is 3.7. The hyperparameter settings in the deep neural network are shown in Table I.

We first observe the change curve of the mean-square error (MSE) loss value of the deep neural network model during training, that is, the square of the difference between the predicted value and the actual value, see Eq. (15). We can observe from Fig. 9 that the network effectively fits the sample data after 10 epochs.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2. \tag{15}$$

In order to observe the prediction effect of our proposed algorithm, we compared the retrieval result of the permittivity and

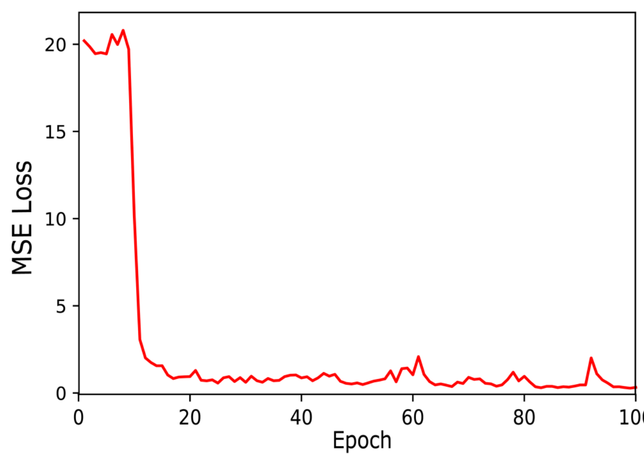

FIG. 9. MSE loss with training epoch.

Published under license by AIP Publishing.

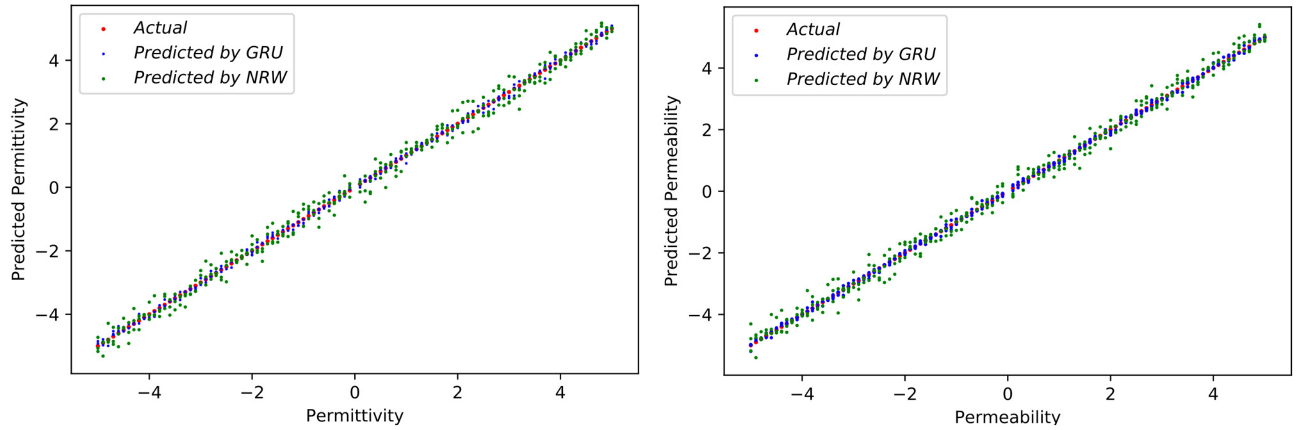

FIG. 10. Comparison of electromagnetic parameter retrieval results. (a) Predicted result of permittivity. (b) Predicted result of permeability.

permeability with the actual values and the results of the NRW method. As described above, our method of obtaining S parameters (Fig. 2) is simpler than the coaxial line in the NRW method (Fig. 1). Therefore, in order to compare with the NRW method, we need to simulate the new S parameters through coaxial lines in CST. In order to show the comparison results more clearly, we prepared 1000 test samples for each method. Figure 10 shows the scatterplots of the retrieval results of the permittivity and permeability. It can be observed that the retrieval results of our model are closer to the actual values than the NRW method.

Furthermore, in order to verify the generalization ability of our model, we tested the material with different thicknesses and evaluated the model effect with MSE. As shown in Table II, our model's accuracy decreases slightly when predicting electromagnetic parameters for material of different thicknesses. The reason is that we did not generate enough samples due to server performance limitations.

TABLE II. MSE of different thickness material.

	Thickness		
Models	1 mm	2 mm	3 mm
GRU	0.14	0.12	0.15
NRW	0.30	0.28	0.28

TABLE III. Comparison of different models.

Models	MSE
Linear regression	16.75
Polynomial regression	9.82
Ridge regression	15.13
NRW	0.28
GRU	0.12

In addition, in order to prove the effectiveness of adopting GRU as the network model, we made comparisons with other common machine learning models based on the same dataset, including linear regression models (that is, single-layer fully connected networks), 2-degree polynomial regression models, and ridge regression. It can be observed from Table III that compared with other machine learning models using the GRU model can greatly improve the accuracy of the model.

IV. CONCLUSION

In this paper, we propose a novel deep learning-based electromagnetic parameter retrieval method that can solve the permittivity and permeability based on the S parameters. This is the first attempt to use deep learning for electromagnetic parameter retrieval. The simulation shows that the proposed scheme has high accuracy and does not have the multivaluedness problem of the traditional NRW method. It provides a novel, simple, and more universal approach for the retrieval method.

ACKNOWLEDGMENTS

This work was supported by the Fundamental Research Funds for the Central Universities (No. CCNU20GF006) and the Foundation for China Scholarship Council (No. 201906770019).

DATA AVAILABILITY

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

REFERENCES

[1]T. C. Choy, *Effective Medium Theory: Principles and Applications* (Oxford University Press, 2015), Vol. 165.
[2]D. R. Smith and J. B. Pendry, "Homogenization of metamaterials by field averaging," J. Opt. Soc. Am. B **23**, 391–403 (2006).

Published under license by AIP Publishing.

[3]G. Lubkowski, R. Schuhmann, and T. Weiland, "Extraction of effective metamaterial parameters by parameter fitting of dispersive models," Microw. Opt. Technol. Lett. **49**, 285–288 (2007).
[4]R. A. Shore and A. D. Yaghjian, "Traveling waves on two- and three-dimensional periodic arrays of lossless scatterers," Radio Sci. **42**, RS6S21 (2007).
[5]X. Chen, T. M. Grzegorczyk, B.-I. Wu, J. Pacheco, and J. A. Kong, "Robust method to retrieve the constitutive effective parameters of metamaterials," Phys. Rev. E **70**, 016608 (2004).
[6]D. Smith, S. Schultz, P. Markoš, and C. Soukoulis, "Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients," Phys. Rev. B **65**, 195104 (2002).
[7]S. O'brien and J. Pendry, "Magnetic activity at infrared frequencies in structured metallic photonic crystals," J. Phys. Condens. Matter **14**, 4035 (2002).
[8]D. Smith, D. Vier, T. Koschny, and C. Soukoulis, "Electromagnetic parameter retrieval from inhomogeneous metamaterials," Phys. Rev. E **71**, 036617 (2005).
[9]H. J. Lee, H. S. Lee, P. S. Ma, and Y. Y. Kim, "Effective material parameter retrieval of anisotropic elastic metamaterials with inherent nonlocality," J. Appl. Phys. **120**, 104902 (2016).
[10]L. Li Hou, J. Y. Chin, X. M. Yang, X. Q. Lin, R. Liu, F. Y. Xu, and T. J. Cui, "Advanced parameter retrievals for metamaterial slabs using an inhomogeneous model," J. Appl. Phys. **103**, 064904 (2008).
[11]A. Nicolson and G. Ross, "Measurement of the intrinsic properties of materials by time-domain techniques," IEEE Trans. Instrum. Meas. **19**, 377–382 (1970).
[12]W. B. Weir, "Automatic measurement of complex dielectric constant and permeability at microwave frequencies," Proc. IEEE **62**, 33–36 (1974).
[13]J. Baker-Jarvis, M. D. Janezic, J. H. Grosvenor, Jr., and R. G. Geyer, "Transmission/reflection and short-circuit line methods for measuring permittivity and permeability," NASA STI/RECON Technical Report No. 93, 1992.
[14]A. N. Vicente, G. M. Dip, and C. Junqueira, "The step by step development of NRW method," in *2011 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC 2011)* (IEEE, 2011), pp. 738–742.
[15]Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).
[16]T. Qiu, X. Shi, J. Wang, Y. Li, S. Qu, Q. Cheng, T. Cui, and S. Sui, "Deep learning: A rapid and efficient route to automatic metasurface design," Adv. Sci. **6**, 1900128 (2019).
[17]J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv:1412.3555 (2014).
[18]S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput. **9**, 1735–1780 (1997).
[19]S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv:1502.03167 (2015).
[20]A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2019), pp. 8024–8035.

Published under license by AIP Publishing.