Deep learning for design metamaterial electromagnetic induction transparent device

Ziming Wei¹, Zeyulin Zhang¹, Wei Huang¹*, Shan Yin¹ and Wentao Zhang¹
¹Guangxi Key Lab. of Optoelectronic Information Processing, School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China.
*Corresponding author’s e-mail: weihuang@guet.edu.cn

Abstract. In this paper, we propose a deep learning model that can be used to reverse design metamaterial electromagnetic induction transparent (EIT) devices. This is a problem that is difficult to achieve with traditional numerical calculation methods. We use the coordinates of six specific points on the EIT transmission spectrum as the input of the neural network, and then the network can predict the structural parameters of the corresponding EIT device. We cite an example to prove that our method can be used to efficiently reverse design the structure of EIT devices. We believe that this method will open up a new way for the structural design of EIT devices and has great potential for expanding the application of terahertz EIT metamaterials.

Keywords: deep learning, neural network, metamaterial, EIT

1. Introduction

Recently, some very new techniques have already applied to Terahertz (THz) research area, such as THz device with graphene [1], THz spoof surface plasmon polariton coupler device with coherent quantum control [2, 3]. Electromagnetically induced transparency (EIT) is a transparent phenomenon in a three-level atomic system that excites two different paths of quantum destructive interference, which can cause the system to weaken or even not absorb light at the resonance frequency. However, the conditions under which the EIT phenomenon occurs in the atomic system are very harsh. Recently, the terahertz electromagnetic induction transparency phenomenon in metamaterials has been well developed and has attracted widespread research attention. The metamaterial-based EIT-like effect not only avoids the harsh conditions of the quantum EIT effect, but also increases the working bandwidth. At the same time, it also extends the EIT effect from the optical band to the microwave and terahertz bands, which has better flexibility and ease of operation.

Because EIT devices are currently widely used, researchers need efficient methods to design EIT devices. At present, if we already know the geometric parameters of the device, it is not difficult to obtain the spectrum by solving the Maxwell equations by the finite difference time domain (FDTD) method. However, due to the high nonlinearity of Maxwell equations and the uncertainty of boundary
conditions, it is still very difficult to reverse design the mechanism of EIT devices through optical response.

Deep learning is currently one of the most popular numerical calculation methods in many fields. It can update the parameters in the neural network through a large number of learning data sets to reduce the loss function to achieve the purpose of predicting results. Deep learning can solve a large number of nonlinear problems, such as target detection[4], natural language processing[5], data mining[6]. It is also widely used in physical systems, such as Quantum states[7, 8], designing nanostructures[9], and others optical systems[10, 11].

Intuitively, deep learning can be used to solve the nonlinear and unconventional boundary conditions in the process of designing EIT devices. In this paper, we propose a deep learning algorithm to design the geometric parameters of the corresponding EIT device through a given EIT transmission spectrum. Our method can design various types of EIT devices corresponding to the transmission spectrum based on the known EIT transmission spectrum. In order to better illustrate the effect of our method, we used the typical EIT device given in the paper[12]. As shown in figure 1, the unit of each metamaterial EIT device consists of a bright mode and a dark mode. The bright mode is a straight metal structure (cut wires, CWs) and the dark mode is two split-ring resonators (SRRs). The change of the EIT device structure has a great influence on its transmission spectrum.

\[ y_i = f(\sum_{i=0}^{l} W_{jk}X_j + B_k) \] (1)

Figure 1. The schematic figure of the metamaterial EIT device.

In this article, we have designed a deep learning model that can predict the geometric parameters of the structure of the EIT device based on its transmission spectrum.

2. Deep Learning Model

Figure 2 shows the working process of the entire model. The entire neural network part is composed of fully connected layers. There are an input layer, an output layer, and several hidden layers in the neural network. The input of the neural network is the transmission spectrum of the EIT device because each transmission spectrum has its corresponding EIT device structure, and when the structure changes, the spectrum will also change significantly. The output in the output layer is the structural geometric parameters of the EIT device. In the learning process of the network, the formula for signal forward propagation is shown in equation (1).
In this equation, $y_i$ is the calculation result of each neuron to be passed to the next layer, $f$ is the activation function of each layer, $W_{ik}$ is the weight between two adjacent layers, $X_i$ is the input of the current neuron, $B_k$ is the threshold of the current neuron.

![Diagram of deep learning model](image)

**Figure 2.** The schematic figure of our deep learning model.

The activation function is the rectified linear unit (ReLU), which is commonly used and has very high efficiency. Its equation is shown in equation (2).

$$\phi(x) = \max(0, x)$$

With ReLU activation function, the convergence speed is faster when calculating gradient descent. Compared with sigmoid and tanh, the derivative of ReLU is easier to find, which is very important. Because the process of back-propagation to update the parameters requires guidance. On the other hand, ReLU can prevent the gradient from disappearing. When the value is too large or too small, the
derivative of sigmoid and tanh is close to 0, while ReLU is an unsaturated activation function without this phenomenon.

The mean square error (MSE) loss function is used in the network optimization process, and its equation is shown in equation (3).

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]  

(3)

In this equation, \(y_i\) is the label value in the data set, \(\hat{y}_i\) is the predicted value of the neural network, \(n\) is the number of samples in the learning process. When training the deep learning model, we used the root mean square prop (RMSprop) optimization. RMSprop is an improvement of the adaptive gradient algorithm (Adagrad). It uses the root mean square as the denominator, which can alleviate the problem of a rapid decline in the learning rate when using Adagrad and reduce fluctuations.

The batch size and learning rate are also very important parameters in the network. The learning rate directly affects the convergence state of the model, and batch size affects the generalization performance of the model. The two are directly related to the numerator and denominator, and they can also affect each other. The learning rate determines the step length of the weight iteration, so it is a very sensitive parameter. Its impact on the model performance is reflected in two aspects. The first is the size of the initial learning rate, and the second is the transformation scheme of the learning rate. The choice of learning rate is a very meaningful work. We have also done some work on the selection of the batch size. Although model performance is not as sensitive to the batch size as the learning rate, the batch size will become a very critical parameter when further improving model performance. A large batch size can reduce training time and improve stability, but it also leads to a decrease in the generalization ability of the model. Therefore, a suitable value of the batch size can ensure accuracy and fast speed. After many experiments, we have determined that the learning rate is 0.001 and the batch size is 500 under the premise of ensuring the speed and accuracy of network optimization.

3. Examples

In order to show the actual effect of the model we designed, we give a detailed example by using the structure shown in figure 1. As shown in figure 2, we set the number of neurons in the input layer and output layer of the neural network to 6 and 5 respectively. The number of hidden layers in the neural network is 7 layers with corresponding to 200, 500, 800, 800, 800, 500, 200 neurons for each hidden layer. The 6 neurons in the input layer represent the coordinates of a maximum point and two minimum points in the transmission spectrum of the EIT device. Because the transmission spectrum corresponding to this structure is all W-shaped, only the three specific points mentioned above can highly confirm the EIT spectrum. Using this method greatly reduces computational complexity and increases learning efficiency. The 5 neurons on the output layer represent the 5 variable parameters of the EIT structure, including the length L of CW, the side length l of SRRs, the gap g of SRRs, the distance between SRRs and CW, and the width w of all-metal structure. In the training process, we used a data set containing 12,000 samples, of which 80% were used as the training set and 20% were used as the test set. Each sample includes the EIT device structure and the coordinates of three points on the corresponding transmission spectrum.

We conducted many experiments for the optimal number of iterations. After 350 iterations, the MSE of the test set is 0.038, and after 400 iterations, the MSE of the test set is 0.034, indicating that 350 iterations is not the optimal value. However, after 450 iterations and 500 iterations, the MSE of the test set is 0.039 and 0.051, respectively. It can be seen that more than 400 iterations will lead to overfitting, so we finally decided to use 400 iterations. We need to spend 1 hour to train the network using i7-CPU.

But after the training is completed, the prediction result of the network can be obtained in less than 1 second. Our deep learning model can learn the samples in the data set well, and can reverse design the structure of the EIT device through the transmission spectrum of the EIT. In order to see the effect of the model more intuitively, we freely selected two sets of geometric parameters and obtained the corresponding EIT spectrum through numerical calculation, as shown by the black solid line in figure 3. Then, we can get the coordinates of three specific points of the EIT transmission spectrum, and input them into our deep learning network. After a very short time of calculation, the network will output a
set of structural parameters. Through this set of parameters, we can obtain the corresponding EIT transmission spectrum, as shown by the red dotted line in figure 3. In order to better demonstrate the robustness of this model, the structure parameters we have selected are beyond the parameter range of the samples in the data set during the network learning process. The spectra in figure 3(a) (b) correspond to the samples with the smallest and largest loss in the test set. Three of five structural parameters corresponding to the transmission spectrum shown in figure 3(c) are out of range, and five structural parameters corresponding to the transmission spectrum shown in figure 3(d) are all out of range.

Figure 3. The effect of deep learning models.

From these examples shown in figure 3, it can be seen that the prediction of the EIT transmission spectrum obtained by our deep learning model is very close to the direct simulation of the EIT transmission spectrum obtained by numerical calculation. Therefore, we can claim that our deep learning model can output the structural parameters of EIT through a given EIT transmission spectrum.

4. Conclusions
To conclude, we propose a deep learning model to efficiently reverse engineer the structure of terahertz EIT devices. The neural network of this model is composed of fully connected layers, and takes the characteristic points on the EIT transmission spectrum as input, and finally the structural parameters of the EIT device can be obtained. This model avoids the problem of solving the highly nonlinear and complicated boundary conditions of Maxwell’s equations in traditional numerical calculation methods. The high efficiency and accuracy of this model make it have broad application prospects in the field of metamaterial EIT. More importantly, this method can be easily extended to other research fields in optics and materials science.

References
[1] Yin S, Shi X, Huang W, Zhang W, Hu F, Qin Z and Xiong X 2019 Two-Bit terahertz encoder realized by graphene-based metamaterials Electron. 8 12
[2] Huang W, Yin S, Zhang W, Wang K, Zhang Y and Han J 2019 Robust and broadband integrated terahertz coupler conducted with adiabatic following New J. Phy. 21 11
[3] Huang W, Qu X, Yin S, Zubair M, Yuan M, Zhang W and Han J 2021 Quantum engineering enables broadband and robust terahertz surface plasmon-polaritons coupler IEEE J. Sel. Top. Quantum Electron. 27 1-7

[4] Li J, Liang X, Li J, Wei Y, Xu T, Feng J and Yan S 2018 Multistage object detection with group recursive learning IEEE Trans. Multimedia 20 1645-55

[5] Wu Y, Yuan M, Dong S, Lin L and Liu Y 2018 Remaining useful life estimation of engineered systems using vanilla LSTM neural networks Neurocomputing 275 167-79

[6] Ghosh K, Amin SA, Gayen S and Jha T 2021 Chemical-informatics approach to COVID-19 drug discovery: Exploration of important fragments and data mining based prediction of some hits from natural origins as main protease (Mpro) inhibitors J. Mol. Struct. 1224 129026

[7] Liu J, Lim K H, Wood K L, Huang W, Guo C and Huang H-L 2019 Hybrid Quantum-Classical Convolutional Neural Networks Preprint 1911.02998

[8] Clark L A, Huang W, Barlow T M and Beige A 2015 Hidden quantum markov models and open quantum systems with instantaneous feedback Interdisciplinary Symposium on Complex Systems pp143-151

[9] Malkiel I, Mrejen M, Nagler A, Ariel U, Wolf L and Suchowski H 2018 Plasmonic nanostructure design and characterization via deep learning Light. Sci. Appl. 7 60

[10] Karanov B, Chagnon M, Thouin F, Eriksson TA, Bulow H, Lavery D, Bayvel P and Schmalen L 2018 End-to-End deep learning of optical fiber communications J. Lightwave Technol. 36 4843-55

[11] Rahmani B, Loterie D, Konstantinou G, Psaltis D and Moser C 2018 Multimode optical fiber transmission with a deep learning network Light Sci. Appl. 7 69

[12] Gu J et al. 2012 Active control of electromagnetically induced transparency analogue in terahertz metamaterials Nat. Commun. 3 1151