

# Individual Prediction of Electric Field Induced by Deep-Brain Magnetic Stimulation With CNN-Transformer

Xiangyang Xu<sup>1</sup>, Bin Deng<sup>1</sup>, Senior Member, IEEE, Jiang Wang<sup>1</sup>, Member, IEEE, and Guosheng Yi<sup>1</sup>

**Abstract**—Deep-brain Magnetic Stimulation (DMS) can improve the symptoms caused by Alzheimer’s disease by inducing rhythmic electric field in the deep brain, and the induced electric field is rhythm-dependent. However, calculating the induced electric field requires building a voxel model of the brain for the stimulated object, which usually takes several hours. In order to obtain the rhythm-dependent electric field induced by DMS in real time, we adopt a CNN-Transformer model to predict it. A data set with a sample size of 7350 is established for the training and testing of the model. 10-fold cross validation is used to determine the optimal hyperparameters for training CNN-Transformer. The combination of 5-layer CNN and 6-layer Transformer is verified as the optimal combination of CNN-Transformer model. The experimental results show that the CNN-Transformer model can complete the prediction in 0.731s (CPU) or 0.042s (GPU), and the overall performance metrics of prediction can reach: MAE = 0.0269, RMSE = 0.0420, MAPE = 4.61% and  $R^2 = 0.9627$ . The prediction performance of the CNN-Transformer model for the hippocampal electric field is better than that of the brain grey matter electric field, and the stimulation rhythm has less influence on the model performance than the coil configuration. Taking the same dataset to train and test the separate CNN model and Transformer model, it is found that CNN-Transformer has better prediction performance than the separate CNN model and Transformer model in the task of predicting electric field induced by DMS.

**Index Terms**—Deep-brain magnetic stimulation, rhythm-dependent, electric field, CNN-transformer.

## I. INTRODUCTION

THE hippocampus is responsible for long-term memory storage, transformation, and orientation. Animal studies have shown that magnetic stimulation of the hippocampus can improve symptoms caused by Alzheimer’s disease and depression [1], [2], [3], [4], [5]. Deep-brain Magnetic Stimulation (DMS) is derived from transcranial magnetic stimulation (TMS), which adopts a pair of large coils to induce a rhythmic electric field deep in the brain to modulate electrical activity of neurons [6], [7], [8]. The intensity of the induced electric field

is important to reflect the effect of DMS. The current loaded by the DMS stimulator generally does not exceed 20A, and its induced electric field intensity in the brain is usually within 5V/m, which is much lower than the intensity of TMS, which generally exceeds 100 V/m [9]. As a result, there are no reports of side effects at present [10].

The head tissues have different conductivity and relative permittivity in the magnetic field environment with different frequencies, which lead to the rhythm-dependent electric field induced by DMS [11]. The anatomical structure of different human brains are different, which lead to the individual difference in the induced brain electric field intensity under the same magnetic field stimulation [9], [12], [13], [14], [15], [16]. In order to obtain the individual effect of DMS, it is generally necessary to establish the brain voxel model and calculate the induced electric field by finite element method (FEM) [17], [18]. However, this process is time-consuming. It usually takes several hours to build a head voxel model and several minutes to perform FEM [19], [20], [21]. Therefore, the effect of DMS on individuals cannot be obtained in real time, which limits its future clinical application. Previous studies have shown that convolutional neural network (CNN) can be used to predict the electric field induced by TMS [22], [23], [24], [25], [26]. The U-net model based on CNN can predict the spatial distribution of the electric field induced by TMS [22]. The electric field induced by TMS is predicted by two factors: the subject’s head MRI and the coil configuration [22], [23], [24]. However, the DMS-induced electric field is rhythm-dependent [10], which leads us to consider not only the subject’s head MRI and coil configuration, but also the stimulation rhythm.

CNN has excellent performance in extracting data features, but it cannot capture remote features due to the limitation of the size of the receptive field [27]. The Transformer model is originally proposed for Natural Language Processing (NLP) [28], [29], [30], which takes advantage of the self-attention mechanism. To adapt to computational vision tasks, Transformer has been improved as a model of Vision Transformer (ViT) [31]. The ViT compensates for the deficiency of CNN by its excellent long-distance feature capture and modeling ability. However, for small datasets, CNN models often perform better than Transformer models [32], [33]. Therefore, recent studies have attempted to combine CNN and Transformer [34], [35], [36], so that the model structure can inherit the advantages of CNN and Transformer, and retain global and local features to the greatest extent. In this

Manuscript received 17 January 2024; revised 8 May 2024; accepted 30 May 2024. Date of publication 3 June 2024; date of current version 12 June 2024. This work was supported by the National Natural Science Foundation of China under Grant 62071324 and Grant 62171311. (Corresponding author: Guosheng Yi.)

The authors are with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: xuxiangyang@tju.edu.cn; dengbin@tju.edu.cn; jiangwang@tju.edu.cn; guoshengyi@tju.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2024.3408902

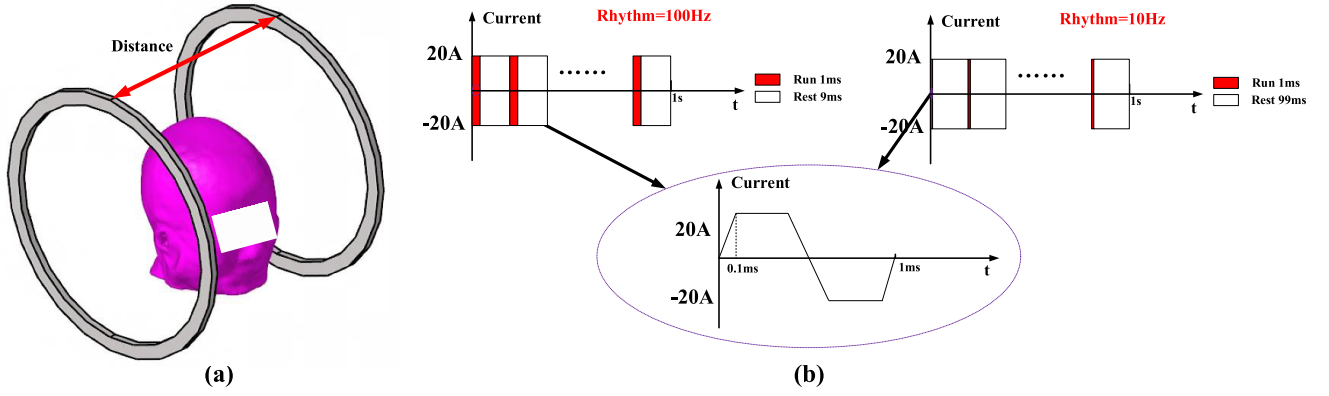


Fig. 1. The stimulation paradigm of DMS. (a) The standard coil configuration of the DMS stimulator. (b) The current waveform loaded in the coil.

paper, a model of CNN-Transformer that combines CNN and Transformer is adopted to obtain induced electric field with a small data set.

We train CNN-Transformer as a predictor to predict rhythm-dependent electric field induced by DMS. First, a dataset is built for training and testing the model, which inputs are MRI, coil configuration and stimulation rhythm, and the outputs are hippocampal electric field and brain grey matter electric field. Then, the best combination of CNN and Transformer layers is obtained and the performance of CNN-Transformer model under this combination is tested. Finally, the prediction performance of CNN-Transformer model, CNN model and Transformer model on rhythm-dependent electric field are compared. Thus, it is verified whether the CNN-Transformer model is suitable for predicting the rhythm-dependent electric field induced by DMS.

## II. MATERIALS AND METHODS

### A. The Stimulation Paradigm of DMS

The standard coil configuration of the DMS stimulator is shown in Fig. 1(a), where two stimulation coils with a diameter of 360mm are placed on the left and right sides of the head, 300mm apart. The current waveform loaded in the coil is shown in Fig. 1(b). The fundamental wave is a bipolar trapezoidal wave with a frequency of 1000Hz and an extreme intensity of  $\pm 20$ A. The rhythm is formed by periodic switching stimulation, and the stimulation rhythm is generally in the range of 10Hz to 100Hz.

### B. Computation Model

The calculation of the rhythm-dependent electric field induced by DMS is based on Maxwell's equations as follows [37]:

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \quad (1)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (2)$$

$$\nabla \cdot \vec{B} = 0 \quad (3)$$

$$\nabla \cdot \vec{D} = \rho \quad (4)$$

where  $\vec{H}$  is the magnetic field intensity,  $\vec{J}$  is the conduction current density,  $\partial \vec{D} / \partial t$  is the displacement current density,  $\vec{E}$

is the electric field intensity,  $\vec{B}$  is the magnetic flux density, and  $\rho$  is the free charge body density. Equation (1) shows that displacement current and conduction current can generate magnetic field. Equation (2) shows that the vortex source of the electric field is the time rate of change of the magnetic flux density. Equation (3) states that the divergence of B is zero, that is, the B line has no beginning and no end. Equation (4) shows that under time-varying conditions, the divergence of D is still equal to the free charge density of the point. In addition, there are three constitutive relations as follows [37]:

$$\vec{D} = \epsilon \vec{E} \quad (5)$$

$$\vec{B} = \mu \vec{H} \quad (6)$$

$$\vec{J} = \sigma \vec{E} \quad (7)$$

where equation (5) shows the relationship between  $\vec{D}$  and  $\vec{E}$  through the relative permittivity  $\epsilon$ . Equation (6) shows the relationship between  $\vec{B}$  and  $\vec{H}$  through the permeability  $\mu$ . Equation (7) shows the relationship between  $\vec{J}$  and  $\vec{E}$  through the conductivity  $\sigma$ . By simplifying and deducing Maxwell's equations, we obtain the electric field E induced by DMS as follows [11] and [12]:

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \nabla \varphi \quad (8)$$

where  $\vec{E}$  is the induced electric field,  $\varphi$  is the scalar potential,  $\vec{A}$  is the magnetic vector potential of the applied magnetic field, and  $\nabla$  is the Hamiltonian operator. The induced electric field  $\vec{E}$  consists of two parts,  $-\partial \vec{A} / \partial t$  is the primary electric field generated by the changing magnetic field, which is determined by the characteristics of the coil and the rate of change of the magnetic field;  $-\nabla \varphi$  is the secondary electric field, induced by the charge in the medium. As the scalar potential  $\varphi$  does not have an analytical solution, it needs to be approximated by FEM. According to Ampere circuital theorem and Gauss's law, the magnetic field partial differential equation and the EF partial differential equation in the electromagnetic field can be derived [37]:

$$\begin{cases} \nabla^2 \vec{A} - \mu \epsilon \frac{\partial^2 \vec{A}}{\partial t^2} + \mu \sigma \vec{E} = 0 \\ \nabla^2 \varphi - \mu \epsilon \frac{\partial^2 \varphi}{\partial t^2} = -\frac{\rho}{\epsilon} \end{cases} \quad (9)$$

where  $\mu$  is the relative permeability of each tissue,  $\epsilon$  is the permittivity of each tissue, and  $\nabla^2$  is the Laplace operator.

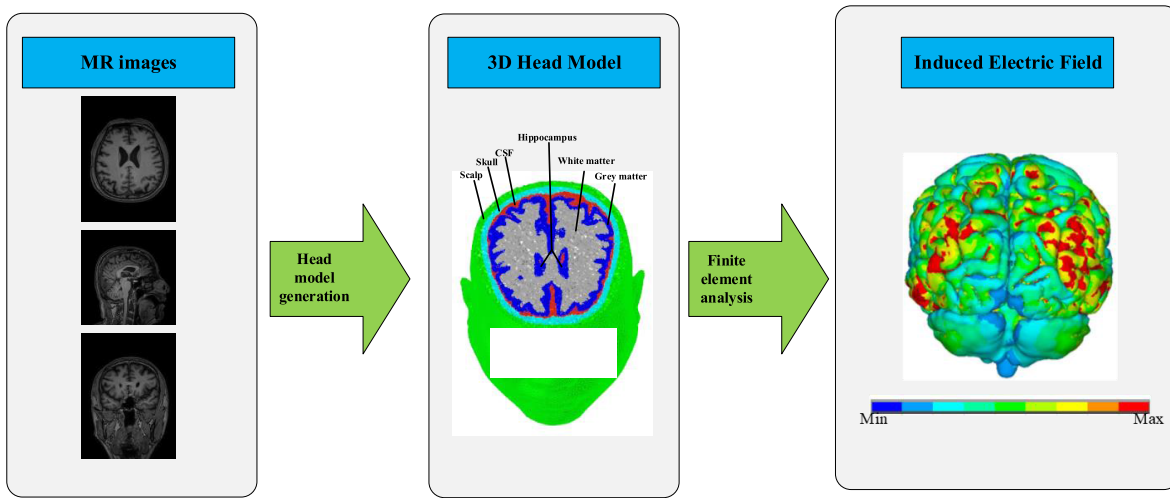


Fig. 2. Procedure for calculating induced electric field.

TABLE I  
THE CONDUCTIVITY AND RELATIVE PERMITTIVITY OF HEAD TISSUES

Rhythm (Hz)	Scalp		Skull		CSF		Grey Matter		White matter	
	Conductivity (S/m)	Relative permittivity	Conductivity (S/m)	Relative permittivity	Conductivity (S/m)	Relative permittivity	Conductivity (S/m)	Relative permittivity	Conductivity (S/m)	Relative permittivity
10	0.20197	25700000	0.0756	10020000	2	109	0.0275	40699000	0.0276	27627000
20	0.20742	24306000	0.0789	4020200	2	109	0.0431	31352000	0.0389	17539000
30	0.21526	22305000	0.08	2020000	2	109	0.0576	22698000	0.0463	11116000
40	0.22423	20021000	0.0805	1196300	2	109	0.0681	16397000	0.0506	7451000
50	0.23329	17719000	0.0807	788910	2	109	0.0752	12107000	0.0533	5289800
60	0.24178	15566000	0.0808	560080	2	109	0.08013	9194500	0.0549	3940300
70	0.2494	13643000	0.0809	419410	2	109	0.0835	7176200	0.0561	3050600
80	0.25606	11973000	0.0810	326970	2	109	0.0859	5740400	0.0570	2436100
90	0.2618	10543000	0.0810	263050	2	109	0.0877	4691500	0.0576	1994900
100	0.26671	9329000	0.0810	217030	2	109	0.0890	3906100	0.0581	1667700

It can be seen from the equations (9) that the induced electric field is not only determined by the applied magnetic field, but also related to the electrical conductivity and relative permittivity of the head tissues.

In this paper, FEM is used to calculate the induced electric field, and the process is shown in Fig. 2: (1) Mimics software (Materialise, Belgium) is used to segment and reconstruct the MRI data of 35 subjects. For each subject, the head structure is divided into six parts: scalp, skull, cerebrospinal fluid (CSF), grey matter, white matter and hippocampus. The conductivity and relative permittivity of the individual head tissues are shown in Table I, with reference to the online database of Italian National Research Council (<http://niremf.ifac.cnr.it/tissprop/>). Hypermesh 2019 software (Altair Engineering, United States) is adopted to mesh the 3D head model, and all the head models are divided into tetrahedral grid with side lengths less than 1mm. (2) ANSYS 2020R2 software (ANSYS, United States) is adopted for finite element calculation of the induced electric field. In the process of finite element calculation, the scalar potential between the scalp, the air and the adjacent head tissue satisfies the Neumann boundary condition [11], [12]:

$$\vec{n} \left( \frac{\partial \vec{A}}{\partial t} + \nabla \varphi \right) = 0 \quad (10)$$

where  $\vec{n}$  is the normal vector.

### C. Data Preparation

The MRI data in this study are obtained from Shanghai Ruijin Hospital. There are 35 subjects aged from 33 to 76, including 21 males and 14 females. Twenty patients with AD and 15 healthy volunteers are included. MRI data from 30 of these individuals are randomly selected for the training set, and MRI data from the remaining 5 individuals are used for the test set.

We select 10 stimulation rhythms (10Hz, 20Hz, 30Hz, 40Hz, 50Hz, 60Hz, 70Hz, 80Hz, 90Hz and 100Hz) in the range of 10Hz to 100Hz at intervals of 10Hz. From 200mm to 300mm, 21 coil spacing of 5mm are selected (200mm, 205mm, 210mm, 215mm, 220mm, 225mm, 230mm, 235mm, 240mm, 245mm, 250mm, 255mm, 260mm, 265mm, 270mm, 275mm, 280mm, 285mm, 290mm, 295mm and 300mm). As a result, a total of 210 combinations of stimulation rhythms and coil configurations are obtained.

We build head models of 35 persons based on MRI data and calculate the electric fields induced by 210 stimulation combinations based on FEM. The information of the dataset is shown in Table II. We obtain all samples by the process shown in Fig. 3(a). The inputs of the dataset are: (1) central slices from coronal view, sagittal view and horizontal view of MRI; (2) coil configuration; (3) stimulation rhythm. The outputs of the constructed dataset are: (4) the maximum

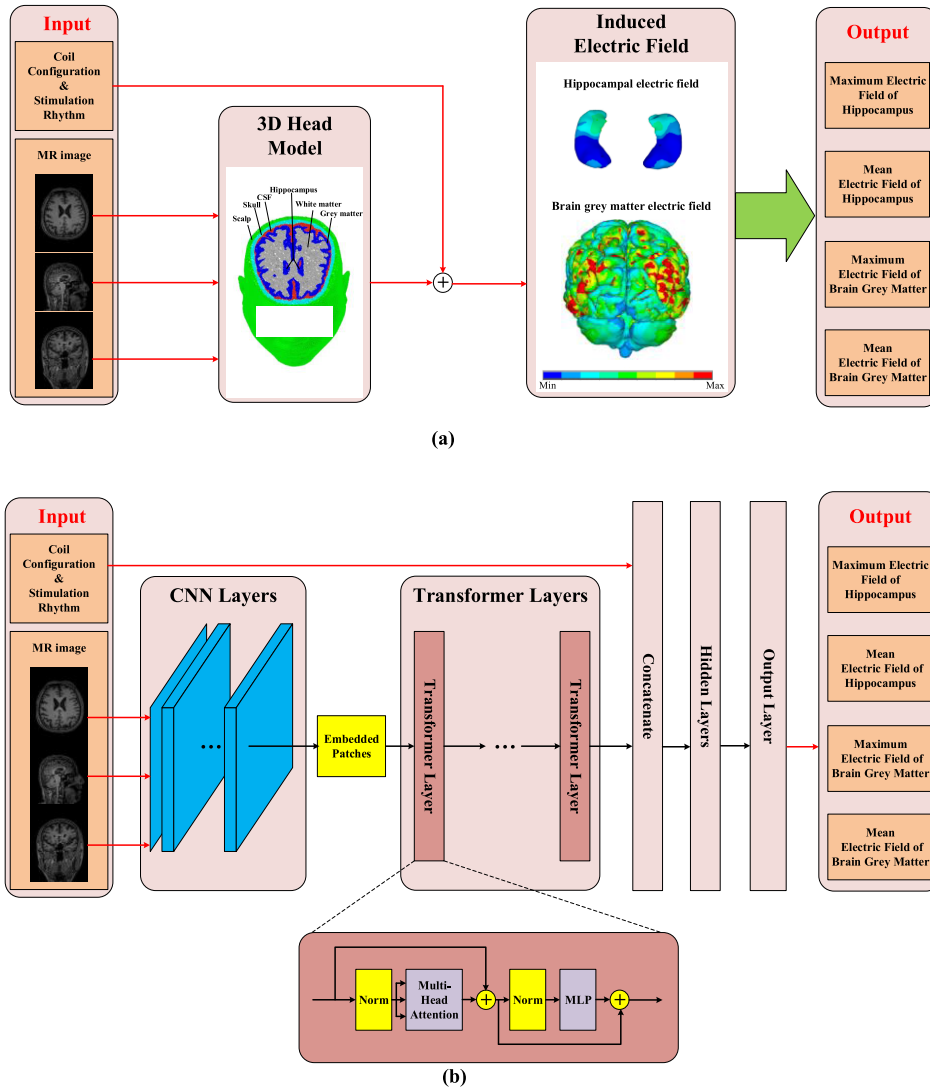


Fig. 3. FEM model and CNN-Transformer model. (a) FEM model. (b) CNN-Transformer model.

TABLE II  
DATASET INFORMATION

Variable type	Variable name	Unit
MR images	Coronal view	-
MR images	Sagittal view	-
MR images	Horizontal view	-
Coil configuration	Distance between coils	mm
Current	Stimulation Rhythm	Hz
Electric Field	Hippocampal EF (Maximum)	V/m
Electric Field	Hippocampal EF (Mean)	V/m
Electric Field	Grey Matter EF (Maximum)	V/m
Electric Field	Grey Matter EF (Mean)	V/m

and mean intensity of the brain grey matter electric field; (5) The maximum intensity and mean intensity of hippocampal electric field. The total number of samples is 7350. Of these, 6300 samples (training set of 30 persons) were used for training and 1050 samples were used for testing (test set of 5 persons).

#### D. CNN-Transformer Model

The framework of the CNN-Transformer model we adopted is shown in Fig. 3(b). It is a hybrid of multi-layer CNN and Transformer. The feature extraction of individual MRI data is the key to predict the electric field induced by DMS for any individual. Referring to the TransUNet model for MRI image segmentation [27], we put the CNN before the Transformer, and actually utilize the CNN as the feature extractor of the Transformer. Treating local and global features separately makes the model design more flexible. The number of layers of CNN and Transformer can be individually adjusted as needed to adapt to predict the electric field of different brain regions.

The part of Transformer adopts the structure of ViT [31]. Since the original Transformer cannot directly process image data, the image needs to be preprocessed. Take image  $x \in \mathbb{R}^{H \times W \times C}$  as an example, where  $H \times W$  is the resolution and  $C$  is the number of channels. First, we split the image  $x$  into  $N$  two-dimensional patches:  $\{x_p^i \in \mathbb{R}^{P \times P \times C} | i = 1, 2, \dots, N\}$ . The size of each patch is  $P \times P$  and the number of patches is  $N = H \times W / (P \times P)$ . Then patch embedding and position encoding are processed. Patch embedding flattens each patch

and makes a linear transformation to map the  $x_p$  to the  $D$ -dimensional vector. At the same time, in order to preserve the spatial position information between the input image patches, a position encoding vector  $E_{position}$  is also needed to be added to the image patch embedding, which is shown as follow:

$$z_0 = [x_{class}; x_p^1 E; x_p^2 E; \cdots; x_p^N E] + E_{position} \quad (11)$$

where  $E \in \mathbf{R}^{(P \cdot P \cdot C) \times D}$  denotes patch embedding and  $E_{position} \in \mathbf{R}^{(N+1) \times D}$  denotes position encoding. In equation (9), a classification vector  $x_{class}$  is added to the vector of length  $N$ , which is used to learn the category information in the training process. The Transformer encoder consists of several layers of multi-head Self-attention (MSA) blocks and multi-layer Perceptron (MLP) blocks. The MSA block of Transformer adopts the Softmax function to calculate the attention weights, and the MLP adopts the GELU activation function. The output of the  $l_{th}$  layer can be expressed as follows:

$$z_l = MLP(LN(MSA(LN(z_{l-1})) + z_{l-1})) + MSA(LN(z_{l-1})) + z_{l-1} \quad (12)$$

where  $LN$  refers to Layer Normalization.  $LN$  is added before each block and residual join is used after each block.

The CNN-Transformer model uses a multi-layer CNN as the feature extractor of the Transformer to generate the feature map of the input. The patch embedding is applied to the patches extracted from the CNN feature map instead of the patches extracted from the original image. Each layer of CNN is composed of three parts: two-dimensional convolution, regularization and RELU activation.

The hidden layer consists of three parts: fully connection, Dropout, and RELU activation, which is able to prevent overfitting. The output layer consists of four neurons and is used to fit four sets of outputs.

For the CNN-Transformer model, we set the inputs as MRI, coil configuration, and stimulation rhythm, and the outputs as the maximum and mean intensity of the hippocampal and brain grey matter electric fields. Firstly, the features of the MRI central slices from three different views are extracted through the CNN layer to generate feature maps. After patch embedding and position encoding, the feature maps are input into the Transformer layer for processing. It is then concatenated with coil configuration and stimulation rhythms to form feature sequences, which are then passed through the hidden and output layers to output the maximum and mean intensity of the hippocampal electric field and the brain grey matter electric field. Both types of data, coil configuration and stimulation rhythm, are single numbers rather than sequences, so using advanced models for feature extraction on them may reduce training efficiency. Referring to the CNN model for predicting the electric field induced by TMS [23], we concatenate the two separate values into the feature map processed by CNN and Transformer to improve the learning efficiency of the model.

## E. Experimental Settings

The model in this study is designed based on Pytorch 1.12.0, the Adam optimizer is adopted to optimize the training, and

the Mean Square Error (MSE) function is used as the loss function.

We adopt four indicators to evaluate the performance of the model in predicting the frequency dependent electric field induced by DMS: mean Absolute error ( $MAE$ ), root mean square error ( $RMSE$ ), mean absolute percentage error ( $MAPE$ ), and coefficient of determination ( $R^2$ ). The four metrics are calculated as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (14)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| 1 - \frac{\hat{y}_i}{y_i} \right| \quad (15)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (16)$$

where  $y_i$  is the intensity of the induced electric field obtained by FEM,  $\hat{y}_i$  is the intensity of the induced electric field predicted by the model, and  $\bar{y}$  is the mean of the intensity of  $n$  sets of electric fields.  $MAE$  and  $RMSE$  can intuitively reflect the error of the results predicted by the model.  $MAE$  reflects the true error, while  $RMSE$  takes the square of the error and then takes the square root, which magnifies the difference of large errors. Therefore,  $RMSE$  can reflect the size of its maximum error.  $MAPE$  is a combination of absolute and relative values, so it can reflect both the degree of deviation from the predicted value and the relative error of the predicted value.  $R^2$  is a statistical measure used to evaluate the goodness of fit of a regression model. The meaning is the proportion of variance explained by the model to the total variance, and the value ranges from 0 to 1. As  $R^2$  gets closer to 1, the model has a better fit to the data, meaning that the independent variable explains more variation in the dependent variable.

In order to keep the performance of the model unbiased, we adopt the method of 10-fold cross-validation. The dataset is equally divided into 10 subsets after shuffling the order. Training is performed 10 times for each combination of hyperparameters. Each time, 9 of the 10 subsets are used as the training set and the remaining one is used as the validation set until all subsets have been used as the validation set. The  $RMSE$  of the validation set is calculated after each training. The mean  $RMSE$  of 10 times is taken as the  $RMSE$  of this hyperparameter combination. Finally, the hyperparameter combination with the minimum  $RMSE$  is selected to train the data of all 10 subsets to obtain the optimal model. The data of the test set is independent of all the data of the training set and the validation set, avoiding the overlap of the data. The test set is used to evaluate the performance of the optimal model to prove its generalization.

Through the 10-fold cross validation method, we determined the training scheme of CNN-Transformer model as follows: the learning rate is set to 0.001, the Batch size is 256, the total number of iterations is 200, and the optimization algorithm

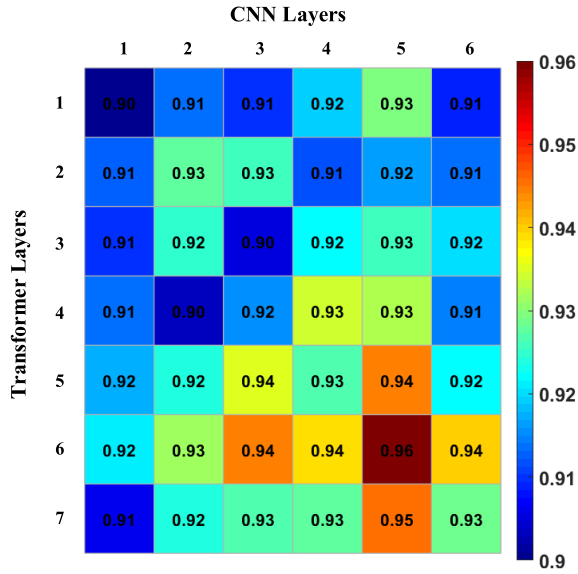


Fig. 4.  $R^2$  with different combinations of CNN layers and transformer layers.

is Adam. The hardware platform we used to perform our experiments is as follows. CPU: Intel(R) Xeon@Silver 4114; GPU: NVIDIA GeForce RTX 3070; Memory: 384GB.

### III. RESULTS AND DISCUSSIONS

#### A. Optimal Combination of CNN and Transformer Layers

In the CNN-Transformer model, the depth of both CNN and Transformer parts has a great impact on the performance of the model. Since deeper networks can learn more abstract and complex features, increasing the depth of the CNN can improve the performance of the model. However, the boost is not linear. As depth increases, the performance gains saturate, and too deep networks may also cause performance degradation. Similar properties hold for the depth of Transformer, increasing depth may improve performance but can also lead to overfitting or training instability. Therefore, choosing the optimal combination of CNN and Transformer layers is very important to improve the prediction performance of this model.

We test CNN layers ranging from 1 to 6 and Transformers layers ranging from 1 to 7, resulting in 42 combinations. Each combination of models is trained for 200 epochs with the training set data, and the  $R^2$  of the model's prediction results on the test set is adopted as the evaluation metric, which are shown in Fig. 4. We find that  $R^2$  tends to decrease after the number of CNN layers exceeds 5 and the number of Transformer layers exceeds 6, and the combination of 5 CNN layers and 6 Transformer layers achieves the largest  $R^2$ . As a result, we confirm that 5-layer CNN and 6-layer Transformer are the optimal combination.

#### B. Prediction Results and Error Analysis

In order to obtain the predictor, we train the CNN-Transformer model of the optimal combination of 5-layer CNN and 6-layer Transformer with the training set. The change of training loss with the number of iterations is shown in the Fig. 5, which loss function of MSE finally converges to the lowest value of 0.0012.

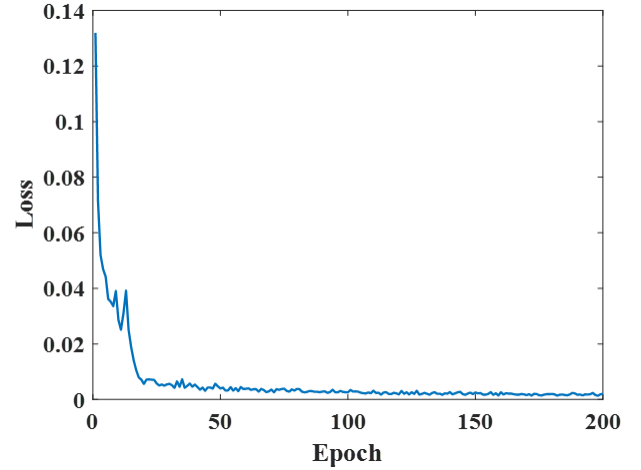


Fig. 5. The change of training loss with the number of epochs.

The comparison of one prediction result of the predictor with the actual data is shown in Table III as an example. For any individual, the maximum and mean values of the induced electric field in the brain grey matter and hippocampus can be predicted by inputting the MRI, coil distance and stimulation rhythm. The results of 35 persons are shown in Fig. 6, where the X-axis represents the calculated value of FEM, the Y-axis represents the predicted value obtained by the model, and the red line represents the reference of  $X=Y$ . The prediction results of the predictor for the brain grey matter electric field and the hippocampal electric field of the training set are compared with the FEM calculation results in Fig. 6(a) and Fig. 6(b), respectively. The prediction results of the predictor for the brain grey matter electric field and the hippocampal electric field of the test set are compared with the FEM calculation results in Fig. 6(c) and Fig. 6(d), respectively.

The  $MAE$ ,  $RMSE$ ,  $MAPE$ , and  $R^2$  of the prediction results of the predictor for the training set and the test set are shown in the Table IV. For both training and test sets, the  $MAE$ ,  $RMSE$  and  $MAPE$  of the hippocampal electric field prediction results are lower than those of the brain grey matter electric field, indicating that the prediction error for the hippocampal electric field is lower. The  $R^2$  of the prediction results for the hippocampal electric field are 0.9773 and 0.9690 in the training set and test set, which indicates that the predictor can predict 97.73% and 96.90% of the variation in the training set and test sets. And  $R^2$  of the prediction results for the brain grey matter electric field are 0.9577 and 0.9402 in the training set and test set, which indicates that the predictor can predict 95.77% and 94.02% of the variation in the training and test sets. The results of  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$  prove that the prediction performance of the predictor for the hippocampal electric field is better than that of the brain grey matter electric field. Comparison of the results on the training and test sets confirms that the model generalizes well. Since the dielectric parameters of the hippocampus are the same as those of the brain grey matter, the difference in prediction performance for brain grey matter and hippocampus mainly depends on their volume differences. The volume of the whole brain grey matter is much larger than that of the hippocampus, which leads to the fact that the number of data features that need to

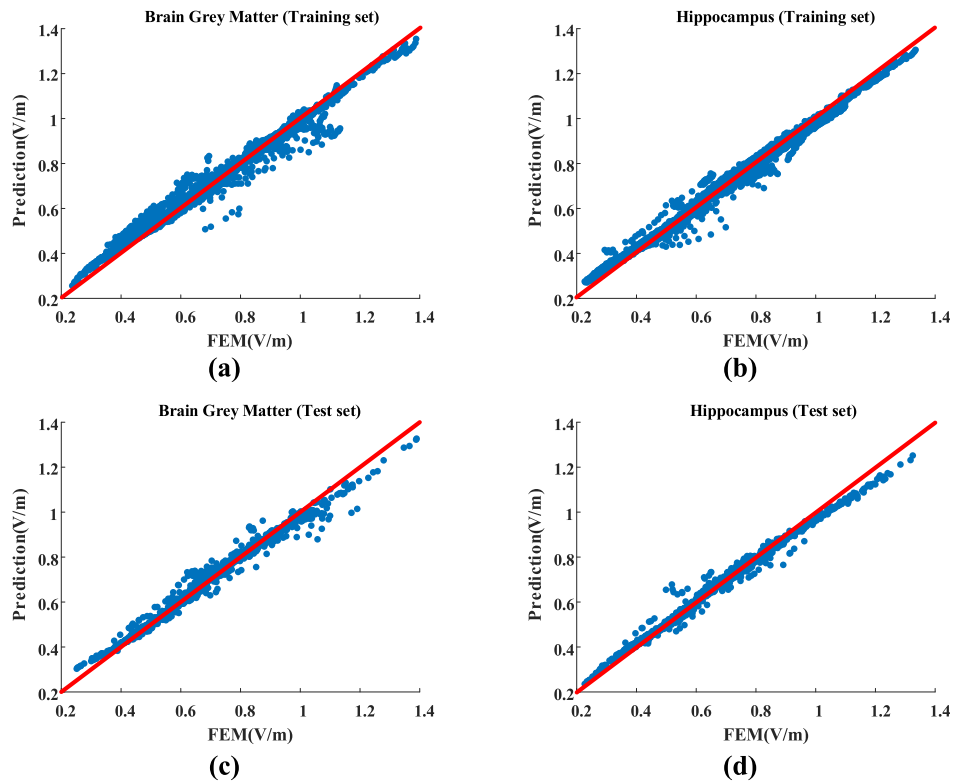


Fig. 6. Comparison of prediction and FEM results. (a) Brain grey matter electric field in training set. (b) Hippocampal electric field in training set. (c) Brain grey matter electric field in test set. (d) Hippocampal electric field in test set.

TABLE III  
COMPARISON OF ACTUAL AND PREDICTED RESULTS

MRI	Coil Distance (mm)	Stimulation Rhythm (Hz)	Prediction (V/m)	Actual (V/m)
	200	15	Grey Matter (maximum): 0.8417	Grey Matter (maximum): 0.8559
			Grey Matter (mean): 0.4604	Grey Matter (mean): 0.4755
			Hippocampus(maximum): 0.5982	Hippocampus(maximum): 0.5851
			Hippocampus(mean): 0.3505	Hippocampus(mean): 0.3464

TABLE IV  
PERFORMANCE COMPARISON ON DIFFERENT DATASETS

Dataset	Tissue	MAE	RMSE	MAPE	R <sup>2</sup>
Training set	Grey Matter	0.0246	0.0469	3.66%	0.9577
Training set	Hippocampus	0.0213	0.0326	3.26%	0.9773
Test set	Grey Matter	0.0291	0.0448	4.94%	0.9402
Test set	Hippocampus	0.0252	0.0385	4.13%	0.9690

TABLE V  
TIME COST COMPARISON

Model	Time Cost
FEM (Including head modeling)	A few hours
CNN-Transformer (CPU)	0.731s
CNN-Transformer (GPU)	0.042s

be considered to predict the hippocampal electrical field alone is much smaller than that of the brain grey matter. Therefore, the prediction performance of hippocampus is better under the same training intensity.

Table V shows the comparison of time cost of FEM and CNN-Transformer model to obtain a set of induced electric fields. In order to calculate the induced electric field of a new individual, FEM needs to first build the head model of the individual, which takes several hours. The CNN-Transformer model takes 0.731s to predict the induced electric field of a

new individual based on CPU, and only 0.042s to predict the induced electric field based on GPU. It indicates that adopting the CNN-Transformer model is able to obtain the rhythm-dependent electric field induced by DMS in close to real time.

### C. Effect of Input Parameters on Model Performance

For the same individual, the stimulation rhythm and coil configuration determine the output of the induced electric field.

In the test set, we test the predictor to predict hippocampal electric field and brain grey matter electric field for all

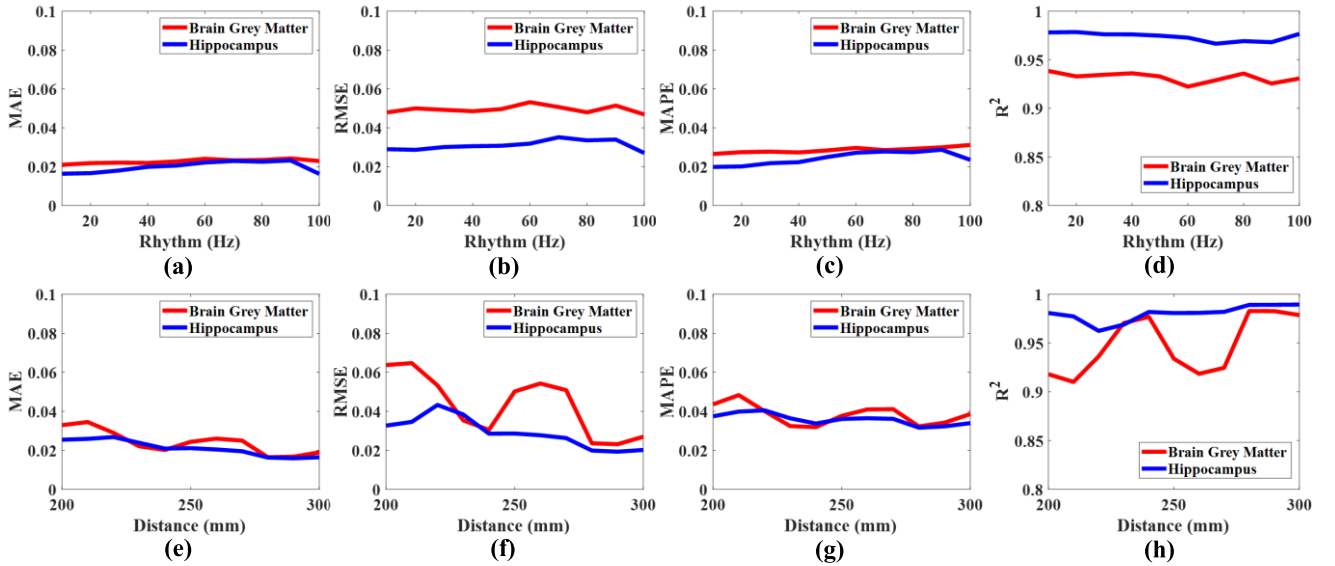


Fig. 7. Effect of input parameters on performance metrics. (a) Effect of stimulation rhythm on  $MAE$ . (b) Effect of stimulation rhythm on  $RMSE$ . (c) Effect of stimulation rhythm on  $MAPE$ . (d) Effect of stimulation rhythm on  $R^2$ . (e) Effect of coil configuration on  $MAE$ . (f) Effect of coil configuration on  $RMSE$ . (g) Effect of coil configuration on  $MAPE$ . (h) Effect of coil configuration on  $R^2$ .

combinations of subject with coil configuration under the same stimulation rhythm. The  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$  are then calculated as metrics of the performance of the predictor in that stimulation rhythm. Then we obtain the performance metrics of all stimulation rhythms, as shown in the Fig. 7(a)-(d). Similarly, in order to obtain the performance metrics of the predictor at different coil configuration, the predictor is tested to predict hippocampal electric fields and brain grey matter electric fields for all combinations of subject with stimulation rhythm at the same coil configuration. The  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$  are then calculated as the performance metrics of the predictor for this coil configuration. The performance metrics of the predictor for all coil configuration are shown in the Fig. 7(e)-(h). Comparing Fig. 7(a)-(h), we find that the performance metrics of  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$  of the predictor all fluctuate significantly with the change of coil configuration, and fluctuate less with the change of stimulus rhythm. The stimulation rhythm, as an input parameter, has significantly less influence on the predictor's performance than the coil configuration. It can be seen from Fig. 7(a)-(d) that as the stimulation rhythm changes, the fluctuation degree of the predictor's performance metrics of the brain grey matter electric field is close to that of the hippocampal electric field. However, it can be seen from Fig. 7(e)-(h), with the change of coil configuration, the fluctuation degree of the predictor's performance metrics of the brain grey matter electric field is far more drastic than that of the hippocampal electric field.

#### D. Effect of Different Tissues on Model Performance

The hippocampus has the same dielectric parameters as the brain grey matter and can be considered as the same type of brain tissue during the calculation of the induced electric field. The difference in prediction performance is likely to depend on the volume difference between the tissues.

In order to verify whether the difference in prediction performance of the model is determined by the volume difference

TABLE VI  
PERFORMANCE COMPARISON OF DIFFERENT TISSUES

Tissues	MAE	RMSE	MAPE	$R^2$
Grey Matter (Whole)	0.0291	0.0448	4.94%	0.9402
Hippocampus	0.0252	0.0385	4.13%	0.9690
Grey Matter (Left)	0.0282	0.0421	4.79%	0.9488
Grey Matter (Right)	0.0280	0.0416	4.72%	0.9494

of brain tissue, we take this model to train the different tissues as grey matter electric field of the left and right hemispheres of the brain separately. By adopting the same hyperparameters for training, we obtained the prediction performance for the grey matter electric field in the left and right hemispheres of the brain compared with the hippocampus and the whole brain grey matter, as shown in Table VI, including  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$ . The prediction performance of the grey matter electric field in the left and right hemispheres of the model brain is very similar, and both of them are between the prediction performance of the whole brain grey matter electric field and the hippocampal electric field. The grey matter volume of the left and right hemispheres is about the same, only half the volume of the whole brain grey matter, and larger than the volume of the hippocampus. It proves that the difference in the prediction performance of different brain regions mainly depends on the difference of volume.

#### E. Performance Comparison With Other Models

Further, we contrast the prediction performance of CNN-Transformer with CNN model and Transformer model for rhythm-dependent electric field induced by DMS. To make a contrast with CNN-Transformer, we set the CNN model to



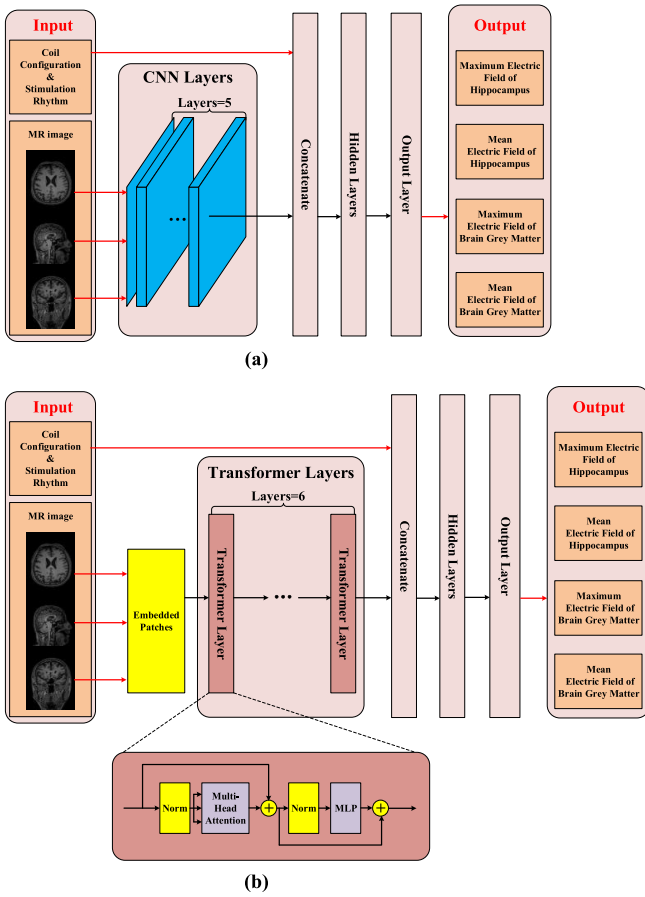


Fig. 8. Comparison of other models. (a) CNN model. (b) Transformer model.

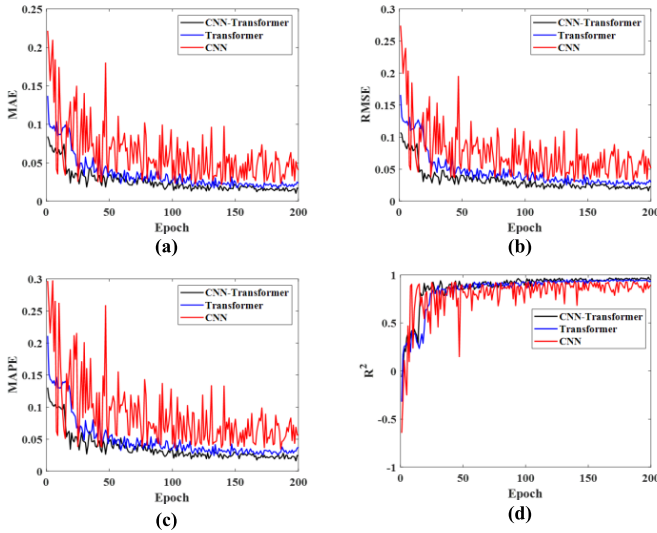


Fig. 9. Comparison of the performance metrics of the models as the number of iterations changes. (a) Comparison of  $MAE$ . (b) Comparison of  $RMSE$ . (c) Comparison of  $MAPE$ . (d) Comparison of  $R^2$ .

5 layers and the Transformer model to 6 layers, which are shown in Fig. 8(a) and Fig. 8(b), respectively.

The above dataset is used to train the CNN model and the Transformer model separately. The hyperparameters used for training are the same as CNN-Transformer: the MSE function

TABLE VII  
PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	MAE	RMSE	MAPE	$R^2$
CNN-Transformer	0.0269	0.0420	4.61%	0.9627
Transformer	0.0415	0.0569	5.37%	0.9381
CNN	0.0421	0.0592	7.20%	0.9192

is used as the loss function of the training, the learning rate is 0.001, the batch size is 256, the number of iterations is 200, and the optimization algorithm is Adam. In the training process, the performance metrics ( $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$ ) of the three models to the training set with the number of iterations are shown in Fig. 9(a)-(d). As can be seen from Fig. 9, the amplitudes of  $MAE$ ,  $RMSE$ ,  $MAPE$ , and  $R^2$  fluctuations with the increase of iterations are all much larger for CNN model than for CNN-Transformer model and Transformer model. This indicates that the CNN model converges slower than CNN-Transformer model and Transformer model in this task. After the number of iterations exceeds 150, all three models gradually tend to converge. We can find from Fig. 9 that for the three metrics of  $MAE$ ,  $RMSE$  and  $MAPE$ : CNN > Transformer > CNN-Transformer. While for the metric of  $R^2$ : CNN-Transformer > Transformer > CNN. It indicates that for the performance of the three models under the training set, CNN-Transformer model is better than Transformer model and CNN model. Then we test the performance of the three models with test set, and the results are shown in the Table VII. Through the comparison of  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$  in the Table VII, we find that the prediction performance for the test set is: CNN-Transformer > Transformer > CNN. It proves that the CNN-Transformer model also has the best generalization performance among the three models in this task.

#### IV. CONCLUSION

In this paper, a CNN-Transformer model is adopted to predict the rhythm-dependent electric field induced by DMS. Compared to FEM, which takes hours to complete head modeling and calculation, CNN-Transformer model only takes 0.731s (CPU) or 0.042s (GPU). It greatly reduces the time cost and approximates being able to obtain the rhythm-dependent electric field induced by DMS in real time.

We build a dataset with a sample size of 7350 for training and testing the CNN-Transformer model. Using 10-fold cross validation, we obtain the optimal combination of hyperparameters: training loss function is MSE, learning rate is 0.001, batch size is 256, number of iterations is 200, and optimization algorithm is Adam. We use the test set to test the different combinations of CNN and Transformer layers, and then we find that the combination of 5-layer CNN and 6-layer Transformer is optimal.

We choose the hippocampal electric field and the brain grey matter electric field as the prediction targets. By comparing the metrics of  $MAE$ ,  $RMSE$ ,  $MAPE$  and  $R^2$ , we find that the CNN-Transformer model has better prediction performance for the hippocampal electric field than the brain grey matter electric field. In addition, we find that the change of stimulation rhythm has a significantly lower impact on

the model performance than the coil configuration. For predicting rhythm-dependent electric field, this property of the CNN-Transformer model is a great advantage.

The CNN-Transformer model is built with the CNN model as the feature extractor of the Transformer model. In order to verify whether the CNN-Transformer model is better than the CNN model or the Transformer model, we take the same dataset to train the CNN model and the Transformer model respectively. For both training and test sets, we find that the CNN-Transformer model performs better than both the CNN model and the Transformer model. It proves that CNN-Transformer is more suitable for predicting the rhythm-dependent electric field induced by DMS.

In future research, the prediction ability of the model for coil configuration should be improved. We will also increase the number of predicted targets to contrast more brain regions.

## REFERENCES

- [1] J. Yuan et al., "Effect of low frequency repetitive magnetic stimulation at shermen (HT7) on sleep quality in patients with chronic insomnia," *Medicine*, vol. 99, no. 30, Jul. 2020, Art. no. e21292.
- [2] N. Sun, Y. He, Z. Wang, W. Zou, and X. Liu, "The effect of repetitive transcranial magnetic stimulation for insomnia: A systematic review and meta-analysis," *Sleep Med.*, vol. 77, pp. 226–237, Jan. 2021.
- [3] L. Xiao et al., "Rhythmic low-field magnetic stimulation may improve depression by increasing brain-derived neurotrophic factor," *CNS Spectrums*, vol. 24, no. 3, pp. 313–321, Jun. 2019.
- [4] J. Zhen et al., "Deep brain magnetic stimulation promotes neurogenesis and restores cholinergic activity in a transgenic mouse model of Alzheimer's disease," *Frontiers Neural Circuits*, vol. 11, p. 48, Jun. 2017.
- [5] S. Sekar, Y. Zhang, H. Miranzadeh Mahabadi, A. Parvizi, and C. Taghibiglou, "Low-field magnetic stimulation restores cognitive and motor functions in the mouse model of repeated traumatic brain injury: Role of cellular prion protein," *J. Neurotrauma*, vol. 36, no. 22, pp. 3103–3114, Nov. 2019.
- [6] Z. Wang et al., "Low field magnetic stimulation promotes myelin repair and cognitive recovery in chronic cuprizone mouse model," *Clin. Experim. Pharmacol. Physiol.*, vol. 48, no. 8, pp. 1090–1102, Aug. 2021.
- [7] J. Zhen et al., "Gamma rhythm low field magnetic stimulation alleviates neuropathologic changes and rescues memory and cognitive impairments in a mouse model of Alzheimer's disease," *Alzheimer's Dementia, Transl. Res. Clin. Intervent.*, vol. 3, no. 4, pp. 487–497, Nov. 2017.
- [8] Z. Sun, T. Jiang, Y. Wu, C. Ma, Y. He, and J. Yang, "Low field magnetic stimulation ameliorates schizophrenia-like behavior and up-regulates Neuregulin-1 expression in a mouse model of cuprizone-induced demyelination," *Frontiers Psychiatry*, vol. 9, p. 675, Dec. 2018.
- [9] J. Gomez-Tames, I. Laakso, and A. Hirata, "Review on biophysical modelling and simulation studies for transcranial magnetic stimulation," *Phys. Med. Biol.*, vol. 65, no. 24, Dec. 2020, Art. no. 24TR03.
- [10] S. Rossi et al., "Safety and recommendations for TMS use in healthy subjects and patient populations, with updates on training, ethical and regulatory issues: Expert guidelines," *Clin. Neurophysiol.*, vol. 132, no. 1, pp. 269–306, Jan. 2021.
- [11] X.-Y. Xu, B. Deng, J. Wang, and G.-S. Yi, "Effect of stimulation frequency on hippocampal electric field induced by deep-brain magnetic stimulation," *AIP Adv.*, vol. 13, no. 1, Jan. 2023, Art. no. 015012.
- [12] A. Opitz, M. Windhoff, R. M. Heidemann, R. Turner, and A. Thielscher, "How the brain tissue shapes the electric field induced by transcranial magnetic stimulation," *NeuroImage*, vol. 58, no. 3, pp. 849–859, Oct. 2011.
- [13] A. M. Janssen, T. F. Oostendorp, and D. F. Stegeman, "The effect of local anatomy on the electric field induced by TMS: Evaluation at 14 different target sites," *Med. Biol. Eng. Comput.*, vol. 52, no. 10, pp. 873–883, Aug. 2014.
- [14] A. Thielscher, A. Opitz, and M. Windhoff, "Impact of the gyral geometry on the electric field induced by transcranial magnetic stimulation," *NeuroImage*, vol. 54, no. 1, pp. 234–243, Jan. 2011.
- [15] A. M. Janssen et al., "The influence of sulcus width on simulated electric fields induced by transcranial magnetic stimulation," *Phys. Med. Biol.*, vol. 58, no. 14, pp. 4881–4896, Jun. 2013.
- [16] A. Nummenmaa, M. Stenroos, R. J. Ilmoniemi, Y. C. Okada, M. S. Hämäläinen, and T. Rajj, "Comparison of spherical and realistically shaped boundary element head models for transcranial magnetic stimulation navigation," *Clin. Neurophysiol.*, vol. 124, no. 10, pp. 1995–2007, Oct. 2013.
- [17] G. B. Saturnino, K. H. Madsen, and A. Thielscher, "Electric field simulations for transcranial brain stimulation using FEM: An efficient implementation and error analysis," *J. Neural Eng.*, vol. 16, no. 6, Nov. 2019, Art. no. 066032.
- [18] J. Gomez-Tames, A. Hamasaka, A. Hirata, I. Laakso, M. Lu, and S. Ueno, "Group-level analysis of induced electric field in deep brain regions by different TMS coils," *Phys. Med. Biol.*, vol. 65, no. 2, Jan. 2020, Art. no. 025007.
- [19] I. Laakso and A. Hirata, "Fast multigrid-based computation of the induced electric field for transcranial magnetic stimulation," *Phys. Med. Biol.*, vol. 57, no. 23, pp. 7753–7765, Nov. 2012.
- [20] S. Aonuma et al., "A high-resolution computational localization method for transcranial magnetic stimulation mapping," *NeuroImage*, vol. 172, pp. 85–93, May 2018.
- [21] Y. Huang, A. Datta, M. Bikson, and L. C. Parra, "Realistic volumetric approach to simulate transcranial electric stimulation—ROAST—A fully automated open-source pipeline," *J. Neural Eng.*, vol. 16, no. 5, Jul. 2019, Art. no. 056006.
- [22] T. Yokota et al., "Real-time estimation of electric fields induced by transcranial magnetic stimulation with deep neural networks," *Brain Stimulation*, vol. 12, no. 6, pp. 1500–1507, Nov. 2019.
- [23] O. F. Afuwape, O. O. Olafasakin, and D. C. Jiles, "Neural network model for estimation of the induced electric field during transcranial magnetic stimulation," *IEEE Trans. Magn.*, vol. 58, no. 2, pp. 1–5, Feb. 2022.
- [24] K. A. Sathi, M. A. Hossain, M. K. Hosain, N. H. Hai, and M. A. Hossain, "A deep neural network model for predicting electric fields induced by transcranial magnetic stimulation coil," *IEEE Access*, vol. 9, pp. 128381–128392, 2021.
- [25] M. Tashli et al., "Prediction of electric fields induced by transcranial magnetic stimulation in the brain using a deep encoder–decoder convolutional neural network," *BioRxiv*, doi: [10.1101/2022.10.27.513583](https://doi.org/10.1101/2022.10.27.513583).
- [26] H. Li, Z.-D. Deng, D. Oathes, and Y. Fan, "Computation of transcranial magnetic stimulation electric fields using self-supervised deep learning," *NeuroImage*, vol. 264, Dec. 2022, Art. no. 119705.
- [27] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," 2021, *arXiv:2102.04306*.
- [28] A. Vaswani et al., "Attention is all you need," in *Proc. NIPS*, 2017, pp. 6000–6010.
- [29] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.
- [30] T. B. Brown et al., "Language models are few-shot learners," in *Proc. NIPS*, 2020, pp. 1877–1901.
- [31] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.
- [32] C. Matsoukas, J. Fredin Haslum, M. Söderberg, and K. Smith, "Is it time to replace CNNs with transformers for medical images?" 2021, *arXiv:2108.09038*.
- [33] M. Raghu et al., "TransFusion: Understanding transfer learning for medical imaging," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 3342–3352.
- [34] A. Srinivas, T.-Y. Lin, N. Parmar, J. Shlens, P. Abbeel, and A. Vaswani, "Bottleneck transformers for visual recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 16519–16529.
- [35] H. Touvron et al., "Training data-efficient image transformers & distillation through attention," in *Proc. ICLR*, 2021, pp. 10347–10357.
- [36] T. Xiao, M. Singh, E. Mintun, T. Darrell, P. Dollár, and R. Girshick, "Early convolutions help transformers see better," 2021, *arXiv:2106.14881*.
- [37] D. Griffiths et al., *Introduction to Electrodynamics*. Cambridge, U.K.: Cambridge Univ. Press, 1999, pp. 326–333.