



OPEN

# Attention-assisted hybrid 1D CNN-BiLSTM model for predicting electric field induced by transcranial magnetic stimulation coil

Khaleda Akhter Sathi<sup>1</sup>, Md Kamal Hosain<sup>2✉</sup>, Md. Azad Hossain<sup>1</sup> & Abbas Z. Kouzani<sup>3</sup>

Deep learning-based models such as deep neural network (DNN) and convolutional neural network (CNN) have recently been established as state-of-the-art for enumerating electric fields from transcranial magnetic stimulation coil. One of the main challenges related to this electric field enumeration is the prediction time and accuracy. Despite the low computational cost, the performance of the existing prediction models for electric field enumeration is quite inefficient. This study proposes a 1D CNN-based bi-directional long short-term memory (BiLSTM) model with an attention mechanism to predict electric field induced by a transcranial magnetic stimulation coil. The model employs three consecutive 1D CNN layers followed by the BiLSTM layer for extracting deep features. After that, the weights of the deep features are redistributed and integrated by the attention mechanism and a fully connected layer is utilized for the prediction. For the prediction purpose, six input features including coil turns of single wing, coil thickness, coil diameter, distance between two wings, distance between head and coil position, and angle between two wings of coil are mapped with the output of the electric field. The performance evaluation is conducted based on four verification metrics (e.g. R2, MSE, MAE, and RMSE) between the simulated data and predicted data. The results indicate that the proposed model outperforms existing DNN and CNN models in predicting the induced electrical field with R2 = 0.9992, MSE = 0.0005, MAE = 0.0188, and RMSE = 0.0228 in the testing stage.

The neurological disorders including major depression, traumatic brain injury, parkinson's disease, and post-traumatic stress are commonly treated with transcranial magnetic stimulation (TMS) therapeutic technique as it provides an effective outcome than other therapeutic techniques<sup>1-4</sup>. The magnetic stimulation technique (i.e., TMS) uses a magnetic coil that is placed on the patient's head. The magnetic coil is activated by applying a high-valued external current pulse with a short period of time<sup>5</sup>. The external applied current pulse produces a magnetic field at the surface of the coil that results in an electric field in the conductive brain tissues of the human head<sup>6</sup>. The induced electric field then forms an axial depolarization in the abnormal neural tissues that are responsible for neural disorders<sup>7</sup>. To activate the targeted abnormal neurons with optimum therapeutic effectiveness, the induced electric field needs to be strong enough at the target side than the other non-targeted region<sup>8,9</sup>. In addition to the strength of the electric field, the depth and focality of the induced electric field are the main concerns in effective TMS treatment.

In TMS therapy, both single coil and assembly coils are used. The single TMS coil has drawbacks associated with the focality and depth of the induced electric field, as it cannot provide an optimum tradeoff of these two parameters. The commercially available single TMS coils including figure of eight, halo, V, circular, H, and double cone coils have been suffering the focality-depth tradeoff problem. All of these single coils except H coil use circular-shaped geometry to produce a focal electric field at the targeted side. Whereas, the H coil generates an electric field with a lower focality stimulation at the target side. For this reason, recent researches

<sup>1</sup>Department of Electronics and Telecommunication Engineering, Chittagong University of Engineering & Technology, Chittagong 4349, Bangladesh. <sup>2</sup>Department of Electronics & Telecommunication Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh. <sup>3</sup>School of Engineering, Deakin University, Geelong, VIC 3216, Australia. ✉email: khosain@ete.ruet.ac.bd

and clinical trials are focused on the assembly-type deep TMS (dTMS) coils to maintain focality-depth tradeoff at the target side<sup>10–13</sup>. Among these dTMS coils, the H7 coil<sup>13</sup> dominates the others in terms of focality-depth tradeoff performance. Despite having the superior focality depth tradeoff performance, there is a difficulty related to coil geometry which results in the likelihood of changing trade-offs with varying coil positions. Moreover, the expense of higher and spreading induced electric field at the superficial cortical region can cause serious headaches. By considering this, a halo-V assembly (HVA) coil is developed<sup>14</sup> to provide optimum therapeutic efficacy by ensuring the finest tradeoff at the target side and reducing the expense of stimulation at the non-targeted superficial cortical region. Furthermore, assembly coil has more design parameters than single coil to achieve desired focality-depth tradeoff. Thereby, for conducting new research and finding a new clinical application of the TMS, the electric field of the assembly coils needs to be enumerated and analyzed with the commercially available electromagnetic (EM) simulation software such as SimNIBS, Sim4life, and COMSOL Multiphysics. This task requires a human head model as a conductive medium and a coil placed on the human head model to determine the induced electric field. The required time to design the new assembly coil and human head model and the simulation time to enumerate induced electric field is huge (e.g. computational cost) that tends to be hours or a day-long<sup>15–19</sup>. Considering this fact, deep learning (DL) based models that are normally dependent on the data are used to enumerate EM field values by reducing the computational cost<sup>20</sup>. Commonly used tasks including segmentation, classification, regression, and clustering are performed in the DL model<sup>21–32</sup>. The electric field estimation is normally based on the regression task that utilizes a deep neural network (DNN) based DL model. In the DL model, a complex non-linear relationship between input and output is developed by adding multiple hidden layers<sup>33</sup>.

Recently, some researches are focused on DL-based electric field enumeration techniques from TMS coils in human tissue models. For instance, Yokota et al.<sup>15</sup> proposed a DL model called as U-Net to determine the electric field from a figure of eight TMS coil by considering one coil parameter such as coil position. The database comprised both input and output parameters as the U-Net regressed the electric field in a supervised learning manner. The output parameter such as electric field was simulated in SimNIBS software where the volume head model was imported from FreeSurfer segmentation software. The FreeSurfer software segmented each part of the brain regions from a brain magnetic resonance image (MRI) to generate a volume head model. After creating the database, it was fed to the 3D U-Net model for training, and finally, the electric field was predicted from test data based on the model training performance. The encoder and decoder-based U-Net model performed electric field prediction with an accuracy of 93%. In another study, Afuwape et al.<sup>16</sup> utilized a convolutional neural network (CNN) based DL model to predict electric fields from sixteen different single-type TMS coils. Initially, the creation of the dataset was performed in Sim4Life software by simulating the different coils on the 3D human head that was generated from T1-weighted MRI. After that, the training of the created dataset was performed on a 3D deep CNN model that regresses the output electric field from the feed input dataset. The performance of the model in predicting electric field was found 92% accurate. Moreover, Sathi et al.<sup>17</sup> proposed an electric field enumeration model based on DNN. The model train and test dataset were created from COMSOL Multiphysics 5.0a software that generated input data from a two-shell human head model and coil geometric structure. In addition, the target data were generated after simulating the model for six coil designing parameters of the HVA coil. The training and testing of prediction model were conducted on 100 datasets. The model electric field prediction accuracy on the testing dataset was found approximately 76%.

In terms of computational time, all existing DL-based models have the capability of computing electric fields in a lower period of time. However, the data creation processes from the segmentation software presented in Refs.<sup>15,16</sup> are quite challenging because it requires high contrast T1-weighted MRI brain image. It is also difficult to produce an actual human head model from low-contrast MRI images. Moreover, the time requirement for the segmentation process is also a problem. Another delimitation of these models is that the electric field estimation accuracy depends on the quality of the MRI scans that result in performance deterioration if the picture quality deteriorates<sup>34</sup>. The utilization of a deep network for both encoding and decoding of data can increase time during training. Furthermore, the use of a single coil parameter such as coil position as the input parameter cannot aid a value to effectively determine the induced electric field. In addition, the prediction accuracy of the electric field of the model developed in Ref.<sup>17</sup> is quite low which may cause incorrect estimation of the new unknown data.

By considering all limitations such as the dependency of image quality in 3D DL network<sup>15,16</sup> and low prediction accuracy in DNN model<sup>17</sup>, a new attention-assisted hybrid 1D CNN-BiLSTM model is proposed in this work for predicting electric field induced by TMS coil. The novelty of this work is to develop a new 1D DL model that use numeric data in electric field enumeration as well as to improve enumeration accuracy than the existing 3D and 1D based state-of-the-art models. In this work, an assembly coil with several designing parameters rather than single parameter i.e., coil position is considered. The model regresses the output electric field from continuous numerical data including different parameters of HVA coil to learn the model from these sample data for final prediction. The created 100 samples dataset is fed to the deep DL model to provide complexity and variation of the dataset and find out the relativity of the input data that are responsible for producing output data of electric field by calculating the attention between each of the input features. Finally, the trained regression model predicts a continuous electric field value from testing data that are unknown to the trained model. The main advantage of the model is associated with the improved prediction accuracy than the recently reported DL-based electric field enumeration model and reduced computational time.

The primary contributions of this work are outlined as follows:

- A hybrid CNN and BiLSTM based framework is proposed for the prediction of induced electric field in HVA TMS coil under low-frequency and high amplitude current pulse conditions.

- An attention mechanism for high accuracy is adopted by redistributing the weights of the deep features extracted by CNN-BiLSTM model.
- The prediction results of CNN-BiLSTM-Attention model with state-of-the-art models are compared. The proposed model enhances performance in both prediction accuracy and computational time complexity.

The remaining part of the paper is organized as follows: In “Methods” section, the dataset creation and architecture of the CNN-BiLSTM-Attention model have been expressed in detail. The results and analysis of model performance in terms of prediction accuracy are described in “Results and performance evaluation” section. Finally, in “Conclusion” section, the concluding remarks is included with future directions briefly.

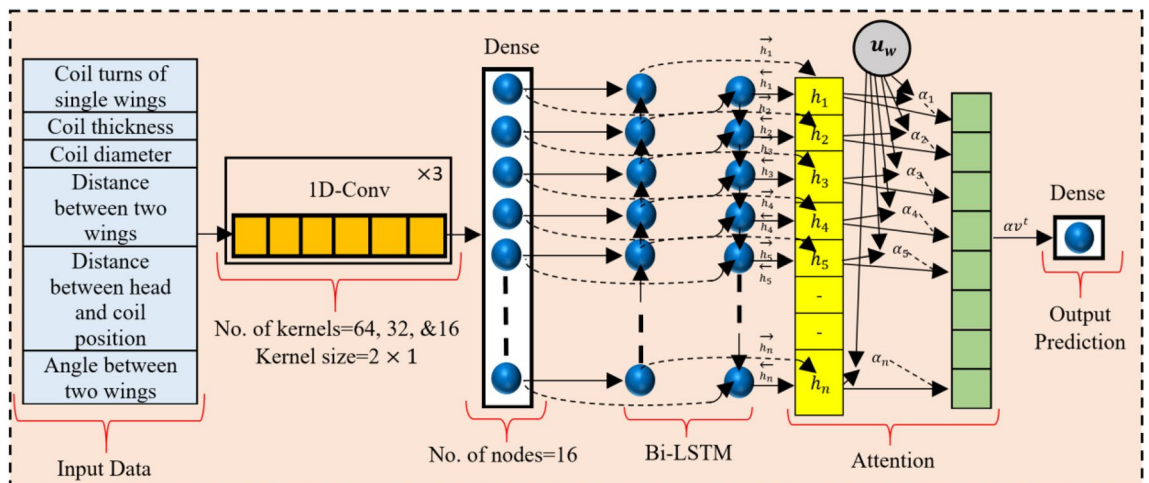
### Methods

The architecture of the proposed attention-based hybrid 1D CNN-BiLSTM model for electric field prediction is shown in Fig. 1. The process of prediction is started by generating a dataset with the composition of input and target features. Here, the input features are the composition of six coil design parameters of the HVA coil and the target feature is the induced electric field. After that, the raw dataset is fed into 1D CNN network for feature extraction by applying three consecutive 1D Conv layers with a kernel size of  $1 \times 2$ . Subsequently, a fully connected layer with 16 nodes is used for the vectorial representation of the features. Then, the sequential model called Bi-LSTM is utilized to learn the valid information from the feature maps by using forward and backward hidden states. Later on, an attention layer is performed that merged important features and chooses the critical features by redistributing the weights. Finally, a fully connected layer is used for electric field prediction. Moreover, the overall processes of the electric field prediction are described below in detail.

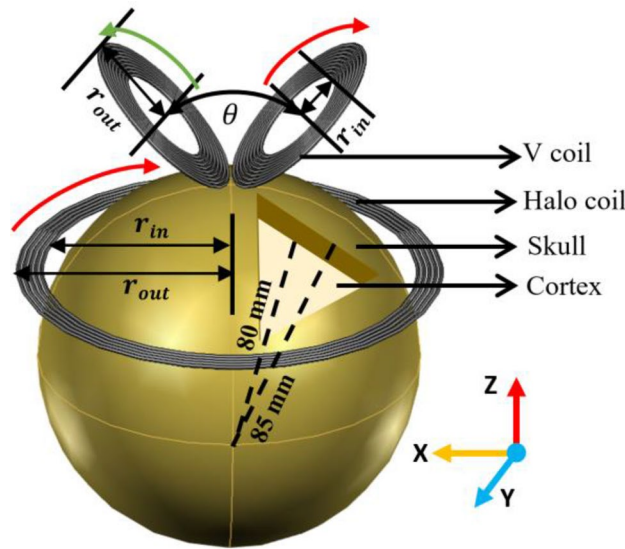
**Configuration of HVA coil with human head model.** A 3D visualization of a HVA coil with two-shell human head model in the cartesian coordinate system is illustrated in Fig. 2. The human head model consists of two layers of tissues including skull and cortex. The radii of the outer and inner layers are 85 mm and 80 mm respectively. On the contrary, the HVA coil is comprised of two single coils such as V-coil and halo coil. The assembly coil has a total of 23 turns among them 5 for the halo coil and 9 for each wings of the V coil. In the HVA coil, the V coil is placed 5 mm above the vertex of the head model and the halo coil is placed 90 mm below the vertex of the head model. All the design parameters of HVA coil are summarized in Table 1. After designing the coil configuration with a torus shape, the copper material with an electrical conductivity of  $5.8 \times 10^7$  S/m is set. In addition, the electromagnetic properties of materials selected for the human head model and coil are listed in Table 2<sup>35,36</sup>.

The electric field in the brain tissues is generated with the HVA coil by applying a current pulse with a high amplitude of 5000A<sup>16</sup> and low-frequency of 2500 Hz. Based on the theory of electromagnetic induction, the varying magnetic field is produced by the rate of change of current in the HVA coil. The electric field<sup>37</sup> is induced in the brain conductive medium by the time-varying magnetic field according to Faraday’s law of induction (1). Equation (1) can be expressed in terms of magnetic vector potential **A**, instead of magnetic field **B** as shown in (2). The vector potential **A** can be expressed as (3): where **r** is the distance from the HVA coil to the field point and **dl** is the vector tangent to the coil at the point.

$$\oint_c \mathbf{E} \cdot d\mathbf{l} = - \int_s \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{s}, \tag{1}$$



**Figure 1.** Overview of the proposed electric field prediction network architecture comprising attention mechanism in the hybrid 1D CNN-BiLSTM model.



**Figure 2.** An illustration of HVA coil configuration with two shell human head model. The clockwise and anti-clockwise directions of coil current are indicated as red and green arrow line respectively.

Coil name		Inner radius $r_{in}$ (mm)	Outer radius $r_{out}$ (mm)	Total coil turns	Angle between two wings $\theta$ (°)
HVA	V	27.5	47.5	18	45°
	Halo	87.5	97.5	5	–

**Table 1.** Design parameters of HVA coil.

Material	Conductivity [S/m]	Relative permittivity	Relative permeability
Skull	0.02	30,380	1
Cortex	4	80	0.99
Copper	$5.8 \times 10^7$	1	0.99

**Table 2.** Electromagnetic properties of HVA coil and two-shell human head model at an operating frequency of 2500 Hz.

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{2}$$

$$\mathbf{A} = \frac{\mu_0 I}{4\pi} \int \frac{1}{r} \cdot d\mathbf{l}. \tag{3}$$

In the neurons, the induced electric field,  $\mathbf{E}$  is the sum of primary electric field component  $\mathbf{E}_1$  and secondary component  $\mathbf{E}_2$ . The  $\mathbf{E}_1$  is responsible for the time-varying magnetic field from HVA coil and  $\mathbf{E}_2$  results from the accumulation of charges on the brain interface as expressed in (4).

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = -\frac{\partial \mathbf{A}}{\partial t} - \nabla V. \tag{4}$$

Using Eq. (3),  $\mathbf{E}$  can be expressed as (5). Where, the minus sign specifies that the direction of the induced electric field is opposite to the direction of the applied current in the HVA stimulating coil. Therefore, the induced electric field is the rate of change of current ( $\frac{dI}{dt}$ ) in the HVA stimulating coil.

$$\mathbf{E} = -\frac{\mu_0}{4\pi} \frac{dI}{dt} \int \frac{1}{r} \cdot d\mathbf{l}. \tag{5}$$

Based on the physics described here, the value of  $\mathbf{E}$  is evaluated on COMSOL Multiphysics 5.0a software. During the simulation, the overall 3D model including the two-shell human head and HVA coil are divided into

several sub-domains with free tetrahedral meshing elements for solving the governing equations which permit accurate numerical results in a low computational time.

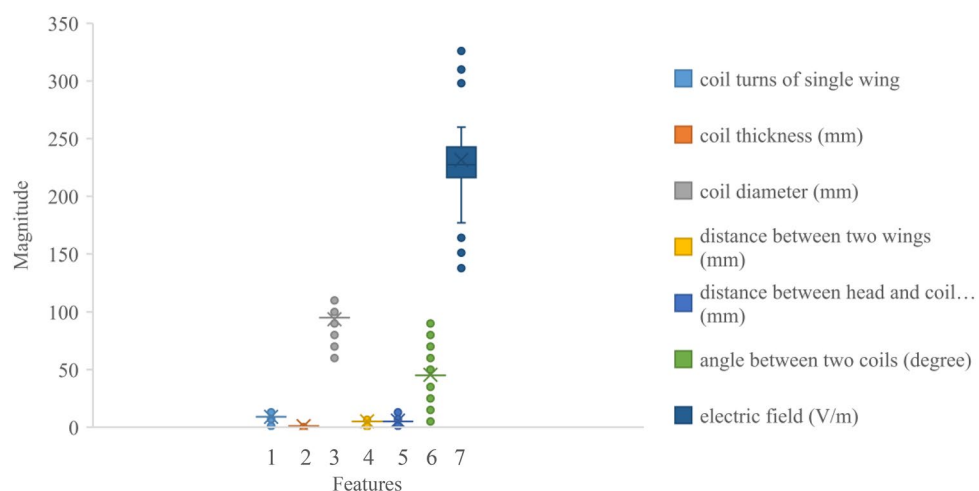
**Dataset creation.** A total of  $N=100$  data samples are collected from COMSOL Multiphysics software and formatted into a .CSV file for processing with the proposed attention-based 1D CNN-BiLSTM model. The  $N$  samples create the dataset, represented as  $D: = \{h, [x_1; x_2; \dots; x_6]_n, E_n\}_{n=1}^N$  with the combination of input  $\{h, [x_1; x_2; \dots; x_6]_n\}_{n=1}^N$  and output  $\{E_n\}_{n=1}^N$  data under the low frequency of 2500 Hz and high amplitude of 5000 A current pulse condition. Where, the output induced electric fields,  $E_n$  are found from six different HVA coil designing parameters including coil turns of single wing ( $x_1$ ), coil thickness ( $x_2$ ), coil diameter ( $x_3$ ), distance between two wings ( $x_4$ ), distance between head and coil position ( $x_5$ ), and angle between two wings ( $x_6$ ) on the human head model. The values of each input parameter and corresponding output variations are summarized in Table 3. Moreover, the box-plot distributions of all input and output parameters are presented in Fig. 3. In this figure, the median values of two input parameters such as coil diameter and angle between two coils are above 50 which designate the dominating factors in calculating the electric field values.

**Proposed attention based 1D CNN-BiLSTM prediction model.** After creating the dataset, the preprocessing technique such as normalization is employed before feeding into the final prediction model. The normalization technique smooths the training time of the network by converting all numerical values in the range of 0 and 1. Then, the normalized data are used for further processing through the proposed prediction model that follows the following sequential steps.

**1D CNN model.** The utilization of 1D CNN model on the normalized continuous numerical data can extract a representative and effective feature by performing a one-dimensional convolution operation with multiple filters. To match the one-dimensional characteristics of the continuous numerical data, the 1D convolutional filters and feature maps are employed with the CNN model. By applying more than one convolutional layers, the 1D CNN model deepens the feature extraction process. Thus a higher level of features makes the prediction task more robust and discriminative. Initially, the preprocessed normalized data with a shape of  $(6 \times 1)$  is used as the input of the 1D CNN model. Then, the input data is passed through three consecutive 1D convolutional (Conv) layers to extract deep features, where each of the Conv layers contains 64, 32, and 16 1D kernels respectively with

Inputs and outputs	Features to form model (unit)	Range
Input	Coil turns of single wing	1–12
	Coil thickness (mm)	0.1–0.9
	Coil diameter (mm)	60–110
	Distance between two wings (mm)	0.5–10
	Distance between head and coil position (mm)	1–15
	Angle between two wings (°)	5–90
Output	Electric field (V/m)	130–300

**Table 3.** Representation of input and output features.



**Figure 3.** The boxplot of original feature visualization.

a size of  $(2 \times 1)$ . After performing the convolution operation, a rectified linear unit (ReLU) activation function is utilized which introduces non-linearity in the model and reduces the overfitting problem as well. The operation of the 1D Conv layer followed by ReLU activation function is performed as follows:

$$y_j^{(k)} = \text{ReLU}_{1,2,3 \in h} \left[ \sum_{i=1}^{N_{(k-1)}} \text{Conv1D}(x_i^{(k-1)}, W_{ij}^{(k)}) + b_j^{(k)} \right], \quad (6)$$

where, the  $i$ th feature map in the  $(k-1)$ th layer is represented as  $x_i^{(k-1)}$  and the  $j$ th feature map in the  $k$ th layer is represented by  $y_j^{(k)}$ . Moreover, the trainable convolutional kernel is denoted as  $W_{ij}^{(k)}$  and  $N_{(k-1)}$  denotes the total number of feature maps in the  $(k-1)$ th layer. The 1D convolution operation without zero-padding is performed on  $x_i^{(k-1)}$  and  $W_{ij}^{(k)}$ . Therefore, the dimension reduction of the feature maps is found at the  $(k-1)$ th layer than the  $k$ th layer. In addition, the bias of the  $j$ th feature map in the  $k$ th layer is represented as  $b_j^{(k)}$ . The operation of ReLU activation function is performed as follows:

$$\text{ReLU}(P_h) = \max(0, P_h), \quad (7)$$

where,

$$P_h = \sum_{i=1}^{N_{(k-1)}} \text{Conv1D}(x_i^{(k-1)}, W_{ij}^{(k)}) + b_j^{(k)}. \quad (8)$$

After passing through all 1D convolutional layers, the obtained 16 feature maps with the size of  $(4 \times 1)$  are fed into one dense layer with 16 nodes. Then, the output features are fed into the Bi-LSTM model for finding valid information from the feature maps.

**Bi-directional LSTM model.** As the dataset is composed of some sort of noisy numerical data, the prediction of the electric field from the noisy data is complex. For estimating electric field and eliminating the noise, the Bi-LSTM network is adopted to the 1D CNN model. The Bi-directional LSTM network is a two-way stacked LSTM network with forward and backward LSTM features. The previous values are learned through the forward LSTM network, and similarly, the future values are learned in the reverse direction by applying a backward LSTM network. The utilization of hidden states in the latter layer help to learn both forward and backward information. The mathematical equation for performing this task through the Bi-LSTM unit is explained as follows:

$$\vec{h}_n = \sigma(W_1 x_t + W_2 \vec{h}_{n-1}) \times \tanh(C_t), \quad (9)$$

$$\overleftarrow{h}_n = \sigma(W_3 x_t + W_4 \overleftarrow{h}_{n-1}) \times \tanh(C'_t), \quad (10)$$

$$h_n^{[p]} = W_5 \vec{h}_n + W_6 \overleftarrow{h}_n, \quad (11)$$

where,  $x_t$  is the input at time  $t$ .  $W$ 's are the weights of gates of LSTM cells.  $\vec{h}_n$  and  $\overleftarrow{h}_n$  are the forward and backward outputs, respectively. The generated Bi-LSTM features from the  $i$ th features in the  $p$ th layer is denoted as  $h_n^{[p]}$  that keeps information in Bi-directional steps. In this paper, the generated feature vectors of 1D CNN and dense are fed to the Bi-LSTM model. An activation function of 'tanh' is adopted to finish the normalization and help to reduce the overfitting problem.

**Attention mechanism.** Generally, all the parameters of the input data do not contribute equally to decide whether the parameter belongs to a particular prediction. Therefore, the utilization of the attention mechanism is performed to emphasize the most important parameters during prediction. In the attention mechanism, the weight is  $\alpha_n$  for each individual Bi-LSTM feature, and  $h_n^{[p]}$  is assigned with a focus on output labels. Mathematically, the followings are computed for the attention function:

$$\mu_i = \tanh(W_d \times \text{flatten}(h_n^{[p]}) + b), \quad (12)$$

$$\alpha_n = \frac{\exp(\mu_i u_w)}{\sum_n \exp(\mu_i u_w)}, \quad (13)$$

$$\alpha v^t = \sum_n \alpha_n h_n^{[p]}. \quad (14)$$

Here,  $h_n^{[p]}$  is the feature vector obtained in the BiLSTM layer, which is passed to a one-layer neural network ( $p=1$ ) to get the  $\mu_i$  as a hidden representation of  $h_n^{[p]}$ . The  $W_d$  and  $b$  are weight matrix and bias vector respectively

that are initialized during the neural network training. The influences of the important parameters can be measured by calculating the similarity between  $\mu_i$  and  $u_w$ . Afterward, a normalized weight,  $\alpha_n$  is obtained by using the softmax function for each input feature. Finally, the attentive feature  $\alpha v^t$  is fed to a dense layer consisting of one neuron. The predicted output,  $E_{Pred}$  can be represented as:

$$E_{Pred} = Linear \left( \sum_j W_{kj} \times \alpha v_j^t + b_k \right), \quad (15)$$

where,  $W_{kj}$  and  $b_k$  are represented as the weight matrix, and bias vector respectively. The activation function 'linear' is added to the proposed model for final electric field prediction. The properties of entire network with their specification and the number of required parameters are demonstrated in Table 4. In this work, the attention-based model with 6641 trainable parameters acquire the best performance for the prediction of the induced electric field.

**Prediction model validation.** To validate the proposed attention-based hybrid 1D CNN-BiLSTM model, the tenfold cross-validation method is utilized that randomly divides the overall data into ten approximately equal sections. Then, each time, one of the splitted sections is selected as the test set, and the rest of the sections are considered as the training set. At every iteration, the model is trained using data shuffling. Finally, the estimation of the model evaluation matrices is performed by taking an average of ten predicted results. The summary of the ten-fold cross-validation methods is depicted as follows:

- Partitioning the raw dataset into ten parts containing an equal number of numerical values in each part.
- One part is selected as the test set for each iteration, and the remaining dataset is utilized as a training set for the model.
- After training, the results of ten iterations are averaged to obtain the final test results.

The selection of optimum hyperparameter values is important to train the proposed prediction model for the superior results. Table 5 represents the optimal values of the hyperparameters of the proposed attention-assisted hybrid 1D CNN-BiLSTM model. The model is compiled using an Adam optimizer with an initial learning rate of 0.01. Moreover, the model training is performed for 100 epochs per fold. The mean absolute error is chosen as a loss function to compute model loss. For conducting training and testing of the proposed model, the Google Colab platform is used with Python version 3.7.13. The model is implemented using Keras = 2.8.0 and TensorFlow = 2.8.2 framework. The Pandas = 1.3.5 and Sklearn = 1.0.2 packages have been used for data preparation and evaluation respectively. During the training, the model occupied 1.43 GB of RAM and 38.58 GB of disk space in the Colab environment.

Layer no.	Layer name	Feature size	Specification	Parameter
1	Input	6 × 1		0
2	1D Convolutional	6 × 64	Filter size: 2 × 1 Filter number: 64	192
3	1D Convolutional	5 × 32	Filter size: 2 × 1 Filter number: 32	4128
4	1D Convolutional	4 × 16	Filter size: 2 × 1 Filter number: 16	1040
5	Dense	4 × 16	–	272
6	Bi-LSTM	4 × 8	Hidden units: 4	672
7	Attention	16	Units: 16	320
8	Dense	1	–	17
				Total = 6641

**Table 4.** Layer properties of proposed attention-based hybrid 1D CNN-BiLSTM model.

Hyper-parameter	Value
Loss function	Mean absolute error
Initial learning rate	0.01
Epochs	100
Optimizer	Adam

**Table 5.** Hyper-parameters setting of the proposed model.

## Results and performance evaluation

The quantitative analysis of the attention-assisted 1D CNN-BiLSTM prediction model is performed based on four verification matrices including coefficient of determination ( $R^2$ ), mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE). The  $R^2$  metric evaluates the quality of regression model by measuring how well the model can predict electric field data from the unknown input data. The preferable value of  $R^2$  in prediction task is considered to be approximately equal to 1. On the other hand, the evaluation metric, MSE is also known as the regression loss function that measures the loss by summing the square of the differences between the predicted value and actual value. Since the loss needs to be lowest, the standard value of MSE for the regression task should be equal to 0. Another loss function called, MAE is measured by averaging all the absolute error values. Similar to the MSE, the model with a MAE value of approximately 0 is preferable. Moreover, the value of RMSE of 0 is desirable for prediction model as it calculates loss by the square root of the sum of square deviation of actual and predicted values over the total number of data,  $n$ . When the approximate value of the RMSE is more close to 0, it means more perfectly fitted the predicted values to the actual values. The mathematical representations of four above-mentioned metrics are expressed in Eqs. (16) to (19). Where, the true output and predicted output are indicated by  $y_{true}$  and  $y_{pred}$  respectively. In addition, the mean value of the ground truth output is represented as  $y_{mean}$  and the total number of data is represented as  $n$ .

$$R^2 = 1 - \frac{\sum_{k=1}^n (y_{true} - y_{pred})^2}{\sum_{k=1}^n (y_{true} - y_{mean})^2} \in [0, 1], \quad (16)$$

where  $y_{mean} = \frac{1}{n} \sum_{k=1}^n y_{true}$ .

$$MSE = \frac{\sum_{k=1}^n (y_{true} - y_{pred})^2}{n}, \quad (17)$$

$$MAE = \frac{\sum_{k=1}^n |y_{true} - y_{pred}|}{n} \in [0, +\infty], \quad (18)$$

$$RMSE = \sqrt{\frac{\sum_{k=1}^n (y_{true} - y_{pred})^2}{n}} \in [0, +\infty]. \quad (19)$$

All the values of performance evaluation matrices are summarized in Table 6. It is clear from Table 6 that the values are almost equal to their preferable values. Based on the preferable values of all the performance matrices, it can be said that the proposed attention-assisted 1D CNN-Bi-LSTM model shows the superiority in prediction of induced electric field with a minor number of errors in the testing data. Moreover, Fig. 4 plots the training process of proposed model in terms of MAE, MSE, and RMSE against 100 iterations (epochs), in which the finest result is achieved after 80 epochs for the evaluation of regression performance.

Figure 5 represents the scatter plot of predicted electric field values against simulated (ground truth) electric field values for the proposed attention-based model. From this graph, it can be observed that all the predicted electric field values are very close to the simulated electric field values that ensure good accuracy between the prediction from attention-based model and simulated data. For this reason, the proposed model is capable of enumerating induced electric field values in a correct manner ranging from 130 to 300 V/m. Moreover, the Kernel Density Estimation (KDE) is illustrated in Fig. 6 to estimate and compare the predicted electric field values with the actual electric field values on the test set in a probability density distribution. According to the performance of the prediction model, the KDE probability density curve of predicted electric field values agrees well with the actual electric field values. Thus, it is recognized that the proposed prediction model attains a superior performance on the test set.

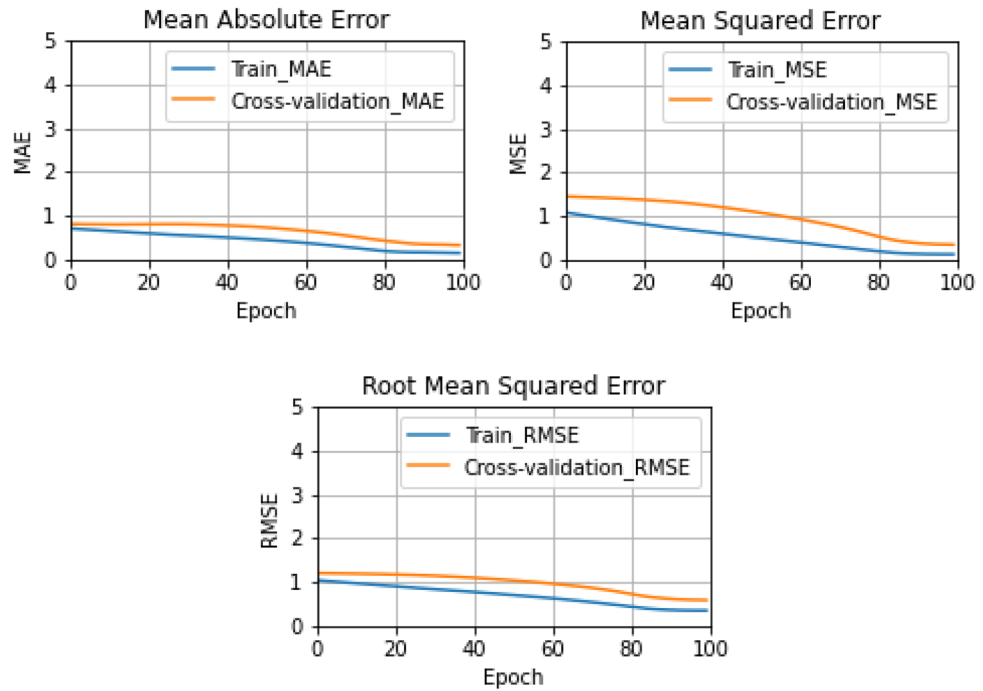
## Discussion

The prediction accuracy,  $R^2$  of different possible regression models such as 1D CNN, BiLSTM, 1D CNN-BiLSTM, 1D CNN with attention, and BiLSTM with attention on validation dataset are illustrated in Fig. 7. Among these DL models, the proposed model provides the highest electric field prediction accuracy of approximately 99.92%. In addition, the model performance is evaluated by comparing the proposed attention-based hybrid model with the state-of-the-art models as presented in Table 7. The works reported in Refs.<sup>15,16</sup> used a single TMS coil type with a single coil parameter of coil position to enumerate the electric field. The main drawback of using a single TMS coil is a lack of tradeoff between focality and depth of stimulation<sup>10</sup>. To ensure an efficient clinical treatment with minimum side effects, the finest tradeoff between focality and depth is needed. In another study<sup>17</sup>, the prediction of electric field from an assembly coil with different coil parameters is presented that meets the

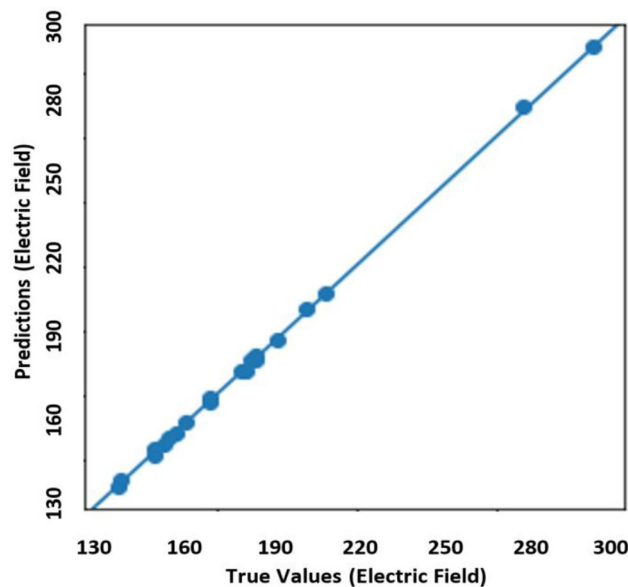
$R^2$	MSE	MAE	RMSE
0.9992	0.0005	0.0188	0.0228

**Table 6.** Performance evaluation matrices for electric field prediction task.



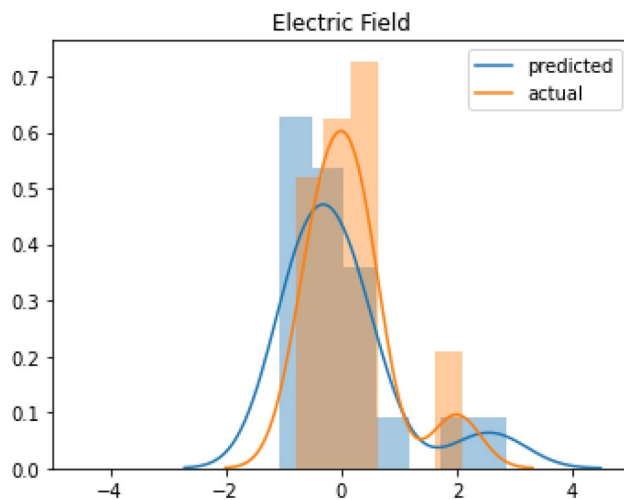


**Figure 4.** Performance evaluation on attention-assisted 1D CNN-BiLSTM model: MAE, MSE, and RMSE.

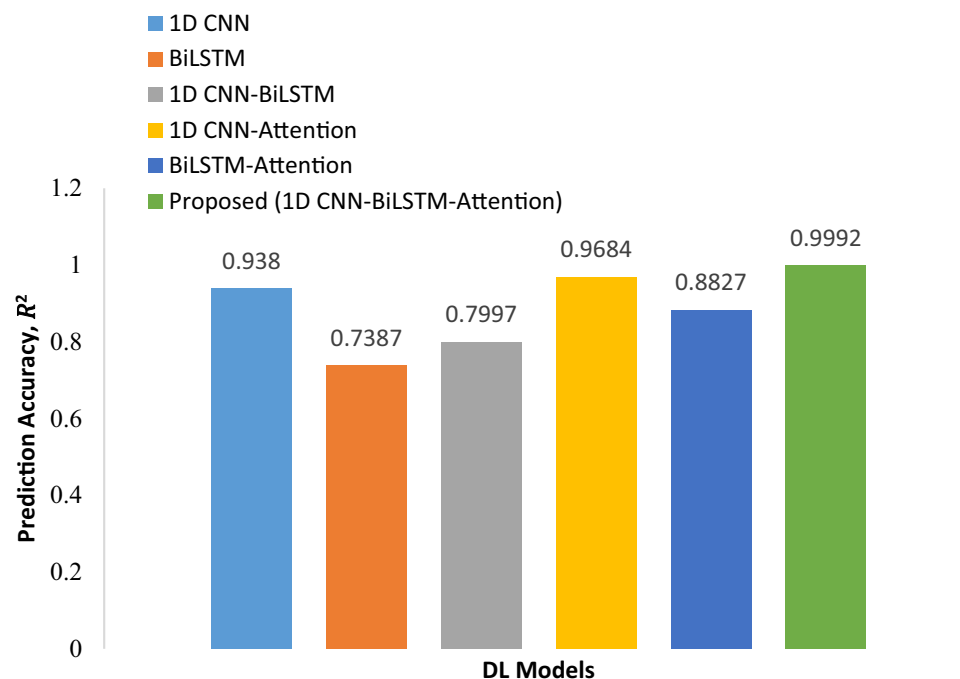


**Figure 5.** Scatter plot of predicted electric field values versus actual data on the test set.

simulation tradeoff between depth and focality. However, the prediction model used in that work provides lower prediction accuracy in terms of  $R^2$  value which is approximately 0.766. Moreover, other evaluation matrices such as MSE, RMSE, and MAE that represent losses of the prediction model are quite larger. For this reason, this research work mainly concerns the improvement of model performance without compromising other required parameters. Compared to the existing works, the proposed attention-based hybrid deep CNN-BiLSTM model ensures optimum prediction accuracy of 0.9992 in terms of  $R^2$  value with the lowest computation time of 0.11 s. Though the database is created based on a spherical human head model rather than a realistic head model, the proposed 1D model prediction performance with a minor variation could ensure optimum therapeutic efficacy including effectiveness and safe stimulation during treatment pertaining to several neurological disorders.



**Figure 6.** KDE probability density curve of the electric field: actual vs. predicted.



**Figure 7.** The prediction accuracy of different models on the validation dataset.

## Conclusion

This work presents a regression model based on an attention-assisted hybrid 1D CNN-BiLSTM network for predicting electric fields induced by HVA TMS coil. Without compromising the computational cost of electric field enumeration, the model improves the prediction accuracy by employing attention layer in the CNN-BiLSTM model. The attention mechanism helps the model to predict electric field values from 130 to 300 V/m by impacting the most relevant parameters of the HVA coil. By feeding a lower-sized database of 100 samples, the model obtained reasonable prediction accuracy of  $R^2 = 0.9992$  that ensures the capability of the model to enumerate electric field from HVA coil with varied coil designing parameters. Without requiring any three-dimensional mathematical model of human head phantom and TMS coil, the proposed method efficiently estimated electric field from a new unknown dataset in a very short time of 0.11 s. Therefore, to meet the requirement of the neurological disorder patients, the proposed attention-based model can aid the TMS manufacturer to design an optimum coil based on the predicted electric fields. In the future, a database with larger samples can be developed by considering several new assembly coils rather than single assembly coils to improve model generalization capability. Moreover, database can be developed based on the numeric data from the realistic anatomical head model to increase the practical feasibility of the work.

Reference	<sup>15</sup>	<sup>16</sup>	<sup>17</sup>	This work
Prediction model	U-Net	Deep CNN	DNN	1D CNN-BiLSTM with Attention Mechanism
Data type	Unstructured: image	Unstructured: image	Structured: numerical value	Structured: numerical value
Number of datasets	261,072	800	100	100
Coil type	Single	Single	Assembly	Assembly
Coil parameter	Coil position	–	Coil position	Coil position
			Coil turns	Coil turns
			Coil thickness	Coil thickness
			Coil diameter	Coil diameter
			Coil angle	Coil angle
			Coil wings distance	Coil wings distance
Prediction task	Segmentation	Regression	Regression	Regression
Computational time	0.541 s	–	0.04 s	0.11 s
Performance evaluation matrix	CC=0.93 PSNR=29 dB MAE=6 RMAD=6%	R <sup>2</sup> =0.92 MAPE=6.2%	R <sup>2</sup> =0.766 MSE=0.184 MAE=0.262 RMSE=0.429	R <sup>2</sup> =0.9992 MSE=0.0005 MAE=0.0188 RMSE=0.0228

**Table 7.** Comparison table of the proposed model with the existing related works.

## Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 14 August 2022; Accepted: 8 February 2023

Published online: 13 February 2023

## References

- Spagnolo, F. *et al.* Bilateral repetitive transcranial magnetic stimulation with the H-coil in parkinson's disease: A randomized, sham-controlled study. *Front. Neurol.* **11**, 1831 (2021).
- Schiena, G., Maggioni, E., Pozzoli, S. & Brambilla, P. Transcranial magnetic stimulation in major depressive disorder: Response modulation and state dependency. *J. Affect. Disord.* **266**, 793–801 (2020).
- BenderPape, T. *et al.* Safety considerations for the use of transcranial magnetic stimulation as treatment for coma recovery in people with severe traumatic brain injury. *J. Head Trauma Rehabil.* **35**(6), 430–438 (2020).
- Isserles, M. *et al.* Deep transcranial magnetic stimulation combined with brief exposure for post-traumatic stress disorder—A prospective multisite randomized trial. *Biol. Psychiatry* **90**, 721 (2021).
- Afuwape, O., Rastogi, P. & Jiles, D. Comparison of the effect of coil configuration and the variability of anatomical structure on transcranial magnetic stimulation. *IEEE Trans. Magn.* **57**, 1–5 (2020).
- Lu, M. & Ueno, S. Comparison of the induced fields using different coil configurations during deep transcranial magnetic stimulation. *PLoS ONE* **12**(6), e0178422 (2017).
- Pashut, T. *et al.* Mechanisms of magnetic stimulation of central nervous system neurons. *PLoS Comput. Biol.* **7**(3), e1002022 (2011).
- Lin, J. C. Transcranial magnetic stimulation therapy for depression and psychiatric disorders [health matters]. *IEEE Microwave Mag.* **17**(8), 23–93 (2016).
- Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin. Neurophysiol.* **120**(12), 2008–2039 (2009).
- de Lara, L. I. N. *et al.* A 3-axis coil design for multichannel TMS arrays. *Neuroimage* **224**, 117355 (2021).
- Lu, M. & Ueno, S. Deep transcranial magnetic stimulation using figure-of-eight and halo coils. *IEEE Trans. Magn.* **51**(11), 1–4 (2015).
- M. Lu, M., Dong, X.-W. & Ueno, S. Deep transcranial magnetic stimulation using deformed halo-circular assembly coil. In *2016 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC)*, Vol. 1, 1103–1105 (IEEE, 2016).
- Gomez-Tames, J. *et al.* Group-level analysis of induced electric field in deep brain regions by different TMS coils. *Phys. Med. Biol.* **65**(2), 025007 (2020).
- Sathi, K. A., Hosain, M. & Hossain, M. Analysis of induced field in the brain tissue by transcranial magnetic stimulation using halo-V assembly coil. *Neurol. Res. Int.* **2022**, 1–10 (2022).
- Yokota, T. *et al.* Real-time estimation of electric fields induced by transcranial magnetic stimulation with deep neural networks. *Brain Stimul.* **12**(6), 1500–1507 (2019).
- Afuwape, O. F., Olafasakin, O. O. & Jiles, D. C. Neural network model for estimation of the induced electric field during transcranial magnetic stimulation. *IEEE Trans. Magn.* **58**, 1–5 (2021).
- Sathi, K. A., Hossain, M. A., Hosain, M. K., Hai, N. H. & Hossain, M. A. A deep neural network model for predicting electric fields induced by transcranial magnetic stimulation coil. *IEEE Access* **9**, 128381–128392 (2021).
- Kybartaitė, A. Computational representation of a realistic head and brain volume conductor model: Electroencephalography simulation and visualization study. *Int. J. Numer. Methods Biomed. Eng.* **28**(11), 1144–1155 (2012).
- Fuchs, M., Wagner, M. & Kastner, J. Development of volume conductor and source models to localize epileptic foci. *J. Clin. Neurophysiol.* **24**(2), 101–119 (2007).
- Rashed, E. A., Gomez-Tames, J. & Hirata, A. Deep learning-based development of personalized human head model with non-uniform conductivity for brain stimulation. *IEEE Trans. Med. Imaging* **39**(7), 2351–2362 (2020).
- Dandil, E. *et al.* Artificial neural network-based classification system for lung nodules on computed tomography scans. In *2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPar)* 382–386 (IEEE, 2014).

22. Shaikat, F. *et al.* Artificial neural network based classification of lung nodules in CT images using intensity, shape and texture features. *J. Ambient. Intell. Humaniz. Comput.* **10**(10), 4135–4149 (2019).
23. Maksimenko, V. A. *et al.* Artificial neural network classification of motor-related eeg: An increase in classification accuracy by reducing signal complexity. *Complexity* **2018**, 1–10 (2018).
24. Cömert, Z. & Kocamaz, A. F. A study of artificial neural network training algorithms for classification of cardiocography signals. *Bitlis Eren Univ. J. Sci. Technol.* **7**(2), 93–103 (2017).
25. Bataineh, M. & Marler, T. Neural network for regression problems with reduced training sets. *Neural Netw.* **95**, 1–9 (2017).
26. Argatov, I. I. & Chai, Y. S. An artificial neural network supported regression model for wear rate. *Tribol. Int.* **138**, 211–214 (2019).
27. Lin, Y., Lu, Q., Tan, B. & Yu, Y. Forecasting energy prices using a novel hybrid model with variational mode decomposition. *Energy* **246**, 123366 (2022).
28. Guresen, E., Kayakutlu, G. & Daim, T. U. Using artificial neural network models in stock market index prediction. *Expert Syst. Appl.* **38**(8), 10389–10397 (2011).
29. Khashei, M. & Bijari, M. Fuzzy artificial neural network (p, d, q) model for incomplete financial time series forecasting. *J. Intell. Fuzzy Syst.* **26**(2), 831–845 (2014).
30. Matos, M., Pinho, S. & Tagarielli, V. Predictions of the electrical conductivity of composites of polymers and carbon nanotubes by an artificial neural network. *Scripta Mater.* **166**, 117–121 (2019).
31. Liang, Y., Lin, Y. & Lu, Q. Forecasting gold price using a novel hybrid model with ICEEMDAN and LSTM-CNN-CBAM. *Expert Syst. Appl.* **206**, 117847 (2022).
32. Lin, Y., Chen, K., Zhang, X., Tan, B. & Lu, Q. Forecasting crude oil futures prices using BiLSTM-attention-CNN model with wavelet transform. *Appl. Soft Comput.* **130**, 109723 (2022).
33. An, Z., Li, S., Wang, J., Xin, Y. & Xu, K. Generalization of deep neural network for bearing fault diagnosis under different working conditions using multiple kernel method. *Neurocomputing* **352**, 42–53 (2019).
34. Sabottke, C. F. & Spieler, B. M. The effect of image resolution on deep learning in radiography. *Radiol. Artif. Intell.* **2**(1), e190015 (2020).
35. Hallez, H. *et al.* Dipole localization errors due to not incorporating compartments with anisotropic conductivities: Simulation study in a spherical head model. *Int. J. Bioelectromagn.* **7**, 134–137 (2005).
36. Gabriel, C. *Complication of the Dielectric Properties of Body Tissues at RF and Microwave Frequencies* (Dept. Phys., King's College London, London, U.K., Brooks Air Force Tech. Rep. AL/OE-TR-1996-0004, 1996).
37. Tsai, Y.-Y. *TMS coil design*. Bachelor thesis. Worcester Polytechnic Institute (2011).

### Author contributions

K.A.S., M.K.H., M.A.H., and A.Z.K. have made substantial contributions to the conception. K.A.S. designed, simulated and wrote the main manuscript text and M.K.H., M.A.H., and A.Z.K. helped to improve the drafted the work. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to M.K.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023