This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

MSFNet: A Multi-Scale Space-Time Frequency Fusion Network for Motor Imagery EEG Classification

Chang Wang[1,2,3,4*], Yang Wu[1,2,3,4], Chen Wang[1,2,3,4], Yaning Ren[1,2,3,4], Jiefen Shen[2,3,4], Ting Pang[2,3,4], Chee Seng Chan[3,5*], Wenjie Ren[3,4*], Yi Yu[2,3*]

[1]The Third Affiliated Hospital of Xinxiang Medical University, Xinxiang, China
[2]College of Medical Engineering, Xinxiang Medical University, Xinxiang, China
[3]Henan International Joint Laboratory of Neural information analysis and Drug intelligent design, Xinxiang, China
[4]Xinxiang Engineering Technology Research Center of Intelligent Medical Imaging Diagnosis, Xinxiang, China
[5]Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding author: Chang Wang (e-mail: wangchang@xxmu.edu.cn).

This work was supported by the Open Project Program of the Third Affiliated Hospital of Xinxiang Medical University (No. KFKTYB202109), Key Scientific Research Project of Henan Higher Education Institutions (No. 23A416001), the major science and technology projects of Henan Province (No. 221100310500).

ABSTRACT Motor imagery electroencephalogram (MI-EEG) classification is essential in brain-computer interface (BCI), and many classification methods have been proposed recently. However, the MI-EEG classification accuracy of the public dataset still has room for improvement, and designing a suitable model to extract and fuse the multi-modality features efficiently is crucial. In this study, we proposed a Multi-scale Space-time Frequency fusion Network (MSFNet) to improve the MI-EEG classification accuracy. The MSFNet comprises data acquisition and preprocessing, multi-scale time-conv fusion unit, multi-scale frequency-conv fusion unit, feature fusion, and classification. Multi-scale time-conv fusion unit can extract multi-scale spatiotemporal features, and multi-scale frequency-conv fusion unit can extract five frequency sub-band features. These two features were concatenated to complete the multi-modality features fusion, and MI-EEG was classified. Average accuracy, kappa value, and F1 score were adopted as the evaluation metric, and BCI Competition 2008 IV 2a and High Gamma datasets were employed to demonstrate the effectiveness of the MSFNet. The superiority of this proposed model was demonstrated by comparison against the state-of-the-art methods, and the classification result is the highest. Overall, we achieved an average accuracy of 80.47%, kappa value of 0.783, and F1 score of 0.743 in the BCI Competition 2008 IV 2a dataset, and we achieved an average accuracy of 93.56%, kappa value of 0.933, and F1 score of 0.915 in High Gamma dataset. This model realized the extraction and efficient fusion of spatiotemporal and frequency domain features and obtained the high-precision MI-EEG classification, which has an important application value.

INDEX TERMS BCI, MSFNet, Multi-scale time-conv fusion unit, Multi-scale frequency-conv fusion unit, MI-EEG classification

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

I. INTRODUCTION

BCI can capture raw neural signals from various cortical regions and modulate the signal into a clinically meaningful output, aiming to restore function to the central and peripheral nervous systems [1,2]. With the development of human-computer interaction, motor imagery EEG (MI-EEG) has been widely studied by researchers. Decoding EEG signals can help people with physical disabilities control external devices. However, MI-EEG signals have small amplitude and low signal-to-noise ratio, and it is difficult to decode MI-EEG [3]. How accurately classifying MI-EEG signals has become the key to research.

Many MI-EEG classification methods based on machine learning (ML) [4,5] and convolutional neural networks (CNN) [6,7] have been proposed recently. Common Spatial Pattern (CSP)[8] can recognize MI-EEG signals by constructing the feature vector with the most significant degree of discrimination, and this method has become a prevalent way to decode MI-EEG signals. Keng et al. used Filter Bank Common Spatial Pattern (FBCSP) algorithm to optimize the subject-specific frequency band of CSP [9]. Hiroshi and Toshihisa used the discriminative filter bank CSP (DFBCSP) to find the frequency bands related to BCI tasks and brain activities [10]. Dornhege et al. used Common Sparse Spectral Spatial Pattern (CSSSP) algorithm to classify MI-EEG signals, and they constructed a complete global Spatio-temporal filter according to the spirit of CSP and typical Spatio-spectral pattern (CSSP) [11]. Malan and Sharma designed a filter bank to divide MI-EEG signals into sub-bands using a dual-tree complex wavelet transform. For ML-based MI-EEG classification methods, they divided MI-EEG signals into the different sub-bands, and used CSP to extract spatial features from each sub-band, and optimized them through a supervised learning framework based on neighborhood component analysis (NCA), trained a support vector machine (SVM) for classification [12]. Graphical features are the newest approaches for identifying underlying patterns of EEG signals, and the phase space reconstruction, Poincare planes, and geometrical features were used in depression and emotion classification [13, 14]. Feature selection plays a vital role in the procession of ML-based MI-EEG classification. In 2019, Yu et al. proposed an automated multivariate empirical wavelet transform (MEWT) algorithm for the different MI-EEG task classifications [15]. In 2021, Sadiq et al. proposed a novel automated framework that reveals the importance of multi-domain features with feature selection for MI-EEG task classification [16]. Although the above methods can achieve good results in the MI-EEG classification, feature extraction and classification are independent. The limitations are that feature selection has a great influence on classification results and that the ML-based MI-EEG classification model has bad generalization performance [17].

Convolutional neural networks have made great strides in MI-EEG classification [18-19]. In 2017, Schirrmeister et al. achieved an average accuracy of 88.90% and a kappa value of 0.852 on HGD using Residual Net [20]. In 2018, Ma et al. proposed a pure RNNs-based parallel method for encoding spatial and temporal sequential raw data with bidirectional Long Short-Term Memory (bi-LSTM) and standard LSTM, respectively [21]. Sakhavi et al. proposed Channel-wise Convolution with Channel Mixing and achieved an average accuracy of 74.46% for the IV-2a Class 4 MI-EEG in the BCI competition [22]. In 2019, Li et al. proposed an MI-EEG classification network named the channel-projection mixed-scale convolutional neural network and achieved an average accuracy of 73.2% and 93.7% on BCI Competition 2008 IV 2a Dataset and High Gamma Dataset, respectively [23]. In 2020, He and Wu proposed EA-CSP-LDA for MI-EEG classification and achieved an average accuracy of 73.53% on MI Dataset 2a [24]. In 2020, Sun et al. proposed a deep learning framework termed SSD-SE-convolutional neural network (CNN) [25]. In 2020, Zhang K et al. proposed five schemes for decoding motor imagery EEG using a deep convolutional neural network (CNN) and achieved the highest subject independence performance with an average accuracy of 84.19%(±9.98%) for two classes of motor imagery [26]. In 2021, Huang JS et al. proposed a classification method using wavelet packet decomposition and deep residual CNN [27]. In 2021, Kumar et al. proposed an EEG monitoring system for epileptical patients and designed a Multi-layered Convolutional Neural Network for the arrangement of eight pneumonic infections [28-29]. We concluded from the previous studies that only spatio-temporal features or frequency features were used for MI-EEG identification, and that the classification accuracy is insufficient. It is generally rare to simultaneously use temporal, spatial, and frequency features. In order to extract different features with high discrimination, we used feature fusion for MI-EEG classification. In 2021, Yang et al. used wavelet convolution and feature fusion methods to achieve an average accuracy of 74.71% and 93.25% on BCI competition IV 2a and High Gamma Dataset, respectively [30]. In 2023, Liu et al. proposed an end-to-end Filter-Bank Multi-scale Convolutional Neural Network (FBMSNet) for MI classification and achieved an average accuracy of 79.17% on BCI competition IV 2a [31]. In 2023, Chen Jiaming et al. designed a filter bank structure with sinc-convolutional layers (FBSCN) for spatio-temporal feature extraction of MI-EEG in four motor rhythms and achieved an average accuracy of 78.20% on BCI competition IV 2a [32]. It proved that using wavelet convolution and feature fusion can effectively classify MI-EEG.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

We concluded from the previous studies that the MI-EEG classification accuracy of the public dataset still has great room for improvement and that it has an important application value in the brain-computer interface. How to design a suitable model to extract the spatiotemporal and frequency features of MI-EEG, and how to fuse the multi-modality features efficiently, which is crucial to the problem of MI-EEG classification. In this paper, we proposed MSFNet for classifying MI-EEG. This model comprises data acquisition and preprocessing, multi-scale time-conv fusion unit, multi-scale frequency-conv fusion unit, feature fusion, and classification. Firstly, we used the data acquisition and preprocessing strategy to remove the noise and artifacts. Secondly, we adopted multi-scale time-conv fusion unit to extract multi-scale spatiotemporal features, and we adopted multi-scale frequency-conv fusion unit to extract five frequency sub-band features. Then, we fused the multi-scale spatiotemporal and frequency features. Finally, we performed the MI-EEG classification. We adopted the average accuracy, kappa value, and F1 score as evaluation metrics to evaluate the model quantitatively. We validated MSFNet on the two public datasets and demonstrated the effectiveness of the proposed model. A comparison against the state-of-the-art methods was analyzed to demonstrate the superiority of our proposed model. We analyzed the effects of different scales and frequency bands in the ablation study.

The main contributions of this paper are as follows:

(1) We proposed MSFNet for MI-EEG classification.

(2) We achieved the most competitive classification results by comparison against the state-of-the-art methods on the two public datasets.

(3) We proved the superiority and robustness of this model by the ablation study.

The remaining sections are organized as follows: Sect. 2 presents the method; Sect. 3 shows experiment results, Sect. 4 discusses this method, and Sect. 5 concludes this study.

II. Method

In this study, we designed a multi-scale space-time frequency fusion network (MSFNet) to complete MI-EEG classification. MSFNet comprises (a) data acquisition and preprocessing, (b) multi-scale time-conv fusion unit, (c) multi-scale frequency-conv fusion unit, (d) feature fusion, and (e) classification, as shown in Fig.1. Multi-scale time-conv fusion unit was designed with different scales, and it can extract the temporal and spatial features of MI-EEG. Multi-scale frequency-conv fusion unit was adopted to extract the different frequency band features of MI-EEG. Feature fusion unit fuses the multi-scale space-time and frequency features. The classification can classify the fusion features and complete MI-EEG classification.

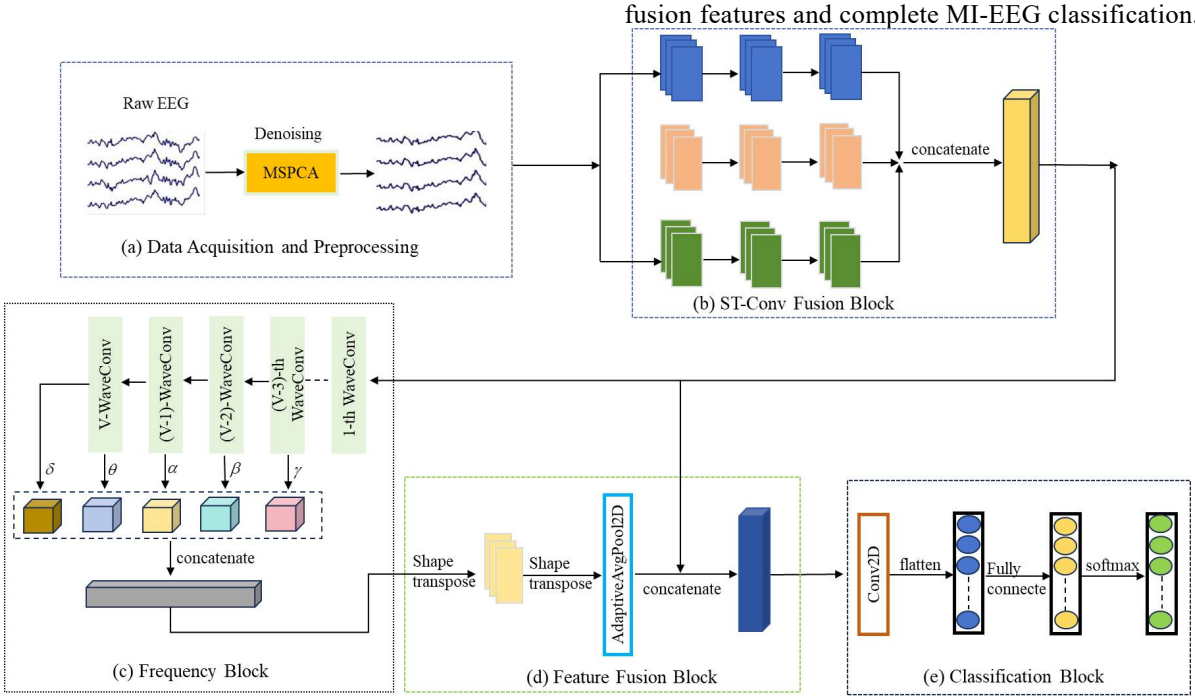


Fig.1. MSFNet model

A. Data Acquisition and Preprocessing

In this study, two public MI-EEG datasets were obtained and preprocessed. BCI Competition 2008 IV 2a Dataset (BCI IV 2a) comprised four classes of motor imagery EEG measurements from nine subjects, namely, left hand, right hand, feet, and tongue. Two sessions, one for training and the other for evaluation, were recorded for each subject. Each session comprised 288 trials of data recorded with 22 EEG channels and 3 monopolar electrooculograms (EOG) channels (with the left mastoid serving as a reference). High Gamma Dataset (HGD) comprised four classes of motor imagery EEG

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

measurements from 14 healthy subjects, namely, left hand, right hand, feet, and rest. MI-EEG was recorded with 128 electrodes with the 500Hz sampling frequency. Each subject had approximately 880 trials of training data and 160 trials of test data. The third-order Butterworth filter was adopted, and the cutoff frequency was set to 0-50Hz.

Motor imagery was cued to start at the moment of 0, and the MI-EEG data was selected from 0.5s before the start of the 0 moments to 4s after the start of the 0 moments. There are 1125 data points in this period. In this study, the size of the trial sample is 22×1125, and 22 represents the channel number, and 1125 represents the data point. The data set is divided as shown in Table 1.

Table 1. MI-EEG dataset

Dataset	Training set	Validation set	Testing set
BCI IV 2a	259	29	288
HGD	792	88	160

MI-EEG signals suffer from several types of noises and artifacts, and they could be described with a linear model. PCA has been widely used as it de-correlates the correlated variables by providing a linear association between different observations. EEG is a non-stationary and nonlinear signal, and wavelet transform was adopted for signal analysis. By combining the benefits of PCA and wavelet, MSPCA can remove noise and artifacts efficiently during MI-ECG preprocessing [33-35].

B. Multi-Scale Time-Conv Fusion Unit

MI-EEG data contains a large of temporal and spatial information. Therefore, we used convolution with the different scales at the beginning of the network to extract relevant features. We designed a multi-scale time-conv fusion unit, which is shown in Fig.1 (b). Convolution layers with different core sizes were adopted to extract the different temporal and spatial features, and these features were fused to enrich the diversity of EEG features. Multi-scale time-conv fusion unit was inspired by the EEGNet model, and it can extract the multi-scale temporal and spatial EEG features, which is shown in Fig. 2. We adopted the backbone of the EEGNet model to extract multi-scale temporal and spatial EEG features. The main parameters of the Multi-scale time-conv fusion unit are shown in Table 2. Where *NF* stands for the number of filters, *KE* stands for kernel size, and Dropout is the ratio of the dropout layer. The standard parameters of *NF*, *KE*, and Dropout were set to 8, (1, 64), and 0.5 respectively in EEGNet [21].

In this study, the parameters of *NF*, *KE*, and Dropout were set in Table 2, and the relationship between *KE* and *NF* in Table 2 is four times. The multi-scale time-conv fusion unit is shown in Figure 3, and the temporal and spatial features with the different scales were concatenated.

C. Multi-Scale Frequency-Conv Fusion Unit

Multi-scale frequency-conv fusion unit is illustrated in Fig.1 (c). We divided the MI-EEG signal into five frequency sub-bands, which are 0-4Hz (δ wave), 4-8Hz (θ wave), 8-12Hz (α wave), 13-30Hz (β wave), and 30-50Hz (γ wave). We used the Daubechies order-4 (Db4) wavelet in this study [36]. Daubechies wavelet can decompose and reconstruct the MI-EEG signal, which can decompose MI-EEG signals into low and high-frequency components. This wavelet has good orthogonality properties and efficient filter implementation [37]. The related conclusion shows that the Db4 wavelet was better at extracting the frequency characteristics of MI-EEG signals. With the implementation of wavelet convolution, the MI-EEG frequency bands can be decomposed into five sub-bands, and the parameters of this unit were set depending on the literature [38].

D. Feature Fusion and Classification

Feature fusion is designed in Fig.1 (d). The five frequency sub-bands obtained from the multi-scale frequency-conv fusion unit were processed by a series of operations of filling with zeros, convolution, batch normalization, and adaptive average pooling with a convolution kernel size of 1x2, and then were concatenated and transformed dimensionally. In the feature fusion, the "squeeze" operation was adopted to aggregate feature maps across time dimensions. Specifically, in order to use the information of channel dependencies, the "squeeze" operation first performs global-average pooling on the input feature maps to squeeze global information, which achieves the purpose of generating channel-wise statistic features. They were finally concatenated with the temporal and spatial features with the three different scales by the multi-scale time-conv fusion unit, and the multi-scale space-time and frequency features were obtained. The classification unit is shown in Fig.1 (e). The output of this unit is the classification probability of the multi-scale space-time and frequency features from the previous feature fusion unit, and it was processed by the convolution layer, flatten layer, fully connected layer, and activation layer. The convolution kernel size in the convolution layer was set to 108x1. The formula for calculating classification probability using the "softmax" function is shown in formula (1):

$$\mathrm{softmax}(z_i) = \frac{e^{z_i}}{\sum_{c=1}^{C} e^{z_c}} \tag{1}$$

where z_i is the output value of the node, and C is the number of categories. The output of multi-classification can be converted into a probability distribution with a range of [0, 1].

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

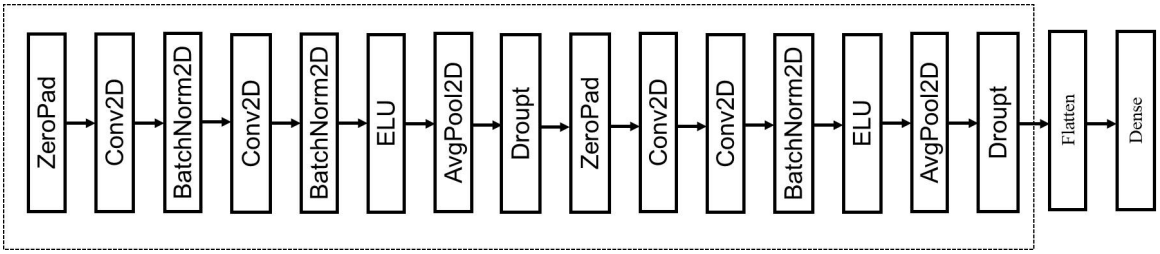


Fig.2. EEGNet model

Table 2. Multi-scale time-conv fusion unit main parameters

Block	Parameter Category	Parameter Value
1st-scale	*NF*	4
	KE	(1,16)
	Dropout	0
2nd-scale	*NF*	8
	KE	(1,32)
	Dropout	0.1
3rd-scale	*NF*	16
	KE	(1,64)
	Dropout	0.2

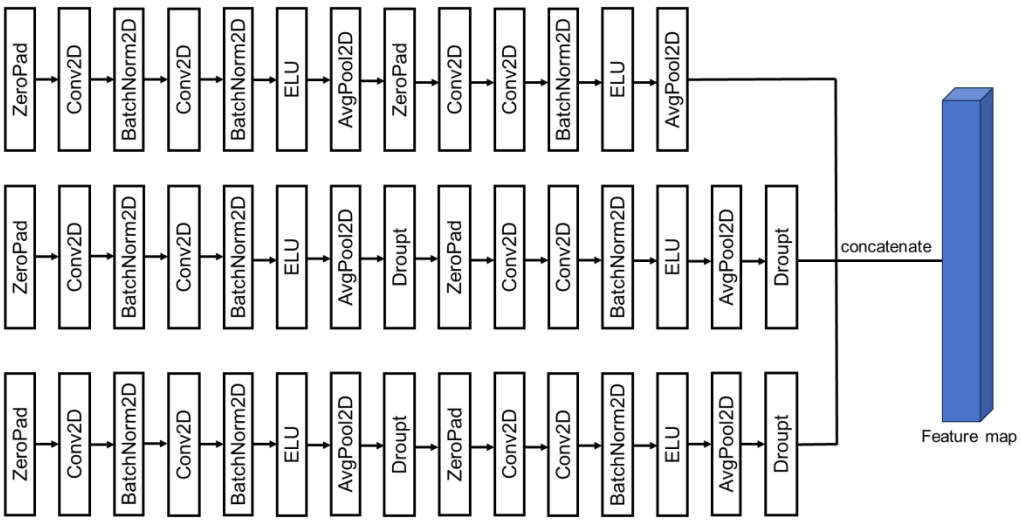


Fig.3. Multi-scale time-conv fusion unit

E. Training details and Evaluation metrics

The experiment environment is PyTorch 1.9 with 4 GPUs of Nvidia Tesla P100 with 16G of memory. The model parameters were set as follows: the epoch is set to 1000, batch size is set to 16, the learning rate is set to 0.001 with weight decay set to 0.1, and cross entropy was adopted as the loss function.

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log(x_{ic}) \qquad (2)$$

where C is the total number of categories and N is the number of samples.

In this study, we adopted classification accuracy, F1 score, and kappa value as evaluation metrics. The classification accuracy can be calculated from the confusion matrix, and the kappa value was calculated as follows:

$$kappa = \frac{Acc - p_e}{1 - p_e} \qquad (3)$$

$$p_e = \frac{\sum_{i=1}^{C} a_i b_i}{n^2} \qquad (4)$$

where a_i is the samples number for each class, b_i is the predicted samples number for each class, C is the categories number, n is the total samples number, and Acc is the overall classification accuracy.

III. Result

Average classification accuracy, F1 score, and kappa value were adopted as evaluation metrics to evaluate this model quantitatively. MSFNet model was tested on two public datasets to demonstrate the effectiveness of this proposed model. The results of the state-of-the-art methods were compared to demonstrate the superiority of this proposed model. The ablation experiment was

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

completed by analyzing the effects of the multi-scale time and frequency bands. t-SNE [40] was adopted to draw the feature distribution to demonstrate the separability of the fusion features.

A. Results of the proposed model

We adopted EEGNet[21], DeepConvNet[20], and ShallowConvNet[20] as the single network models to demonstrate the effectiveness of MSFNet model. EEGNet is a compressed convolutional neural network. The convolutional layer, depth-wise convolutional layer, and separable convolutional layer are the three main convolutional layers, and the SoftMax function as the activation function is used in the last dense layer. DeepConvNet is a more complex network than EEGNet, which contains five convolutional layers. The Elu function is the activation function of the first four layers, and the SoftMax function is the activation function of the last dense layer. ShallowConvNet is the simplest model compared with the other two models, and it only has two convolutional layers, the Square, Clip, and Log functions are defined as the activation function in ShallowConvNet model.

In this study, we adopted MSPCA as the preprocessing strategy, and Fig. 4 is the preprocessing results. Fig. 4(a) is the time-domain signal of the raw MI-EEG, and Fig. 4(b) is the preprocessing results of the raw MI-EEG by MSPCA, and it can be shown that MSPCA can remove the noise and artifacts.

We adopted the five-fold cross-validation strategy to ensure the generalization capability of the proposed model and obtained more reliable classification results. We tested subject A01 in BCI IV 2a by the five-fold cross-validation strategy, and the classification result is shown in Table 3. The cross-validation result is that the average accuracy, F1 score, and kappa value are 83.68%, 0.832, and 0.791, respectively.

Tables 4-5 show the evaluation results of the MSFNet model compared with the above three models on the two public datasets. The evaluation metrics of the MSFNet model are the highest. It can be seen from Table 4 that the average accuracy, F1 score, and kappa value using the MSFNet model are 80.47%, 0.783, and 0.743, respectively. It also can be seen from Table 5 that the average accuracy, F1 score, and kappa value using the MSFNet model are 93.56%, 0.933, and 0.915, respectively. MSFNet can achieve the best results on the two public datasets.

The number of the different model parameters is shown in Table 6. We concluded that EEGNet has less than 10 thousand parameters, that DeepConvNet has hundreds of thousands of parameters, and that MSFNet has more than 10 thousand parameters. MSFNet is simpler than DeepConvNet and ShallowConvNet, and more complex than EEGNet. The training time of this proposed model is 3.5 and 5.7 hours in BCI IV 2a and HGD.

B. Comparison with the state-of-the-art methods

We compared the results of MSFNet with those from other literature to demonstrate the superiority of MSFNet model, which is shown in Table 7. The methods proposed in the literature [23], [30], and [31] can extract the temporal and spatial features, but they cannot extract the frequency features of MI-EEG. But frequency features are an important part of signal characteristics. The method proposed in the literature [38], [39], [40] can extract the temporal, spatial, and frequency features of MI-EEG, but the obtained classification results are not the highest.

We concluded that the average accuracy of MSFNet is higher than the literature [20] by 6.01% and higher than the literature [23] by 7.27% on BCI IV 2a. The average accuracy of MSFNet is higher than the literature [38] by 0.31% on HGD. The average accuracy and kappa value of MSFNet are the highest on the two public datasets compared with the state-of-the-art methods.

Literature [41] [33] achieved 98.3% and 90% average accuracies on BCI III IVa, respectively. BCI III IVa is a public MI-EEG dataset provided in 2003. It contains 118 EEG channels, including two classes (right hand, foot) from 5 subjects, which is different from our study. BCI IV 2a is a more difficult four-category classification problem, and the classification accuracy cannot be comparable with BCI III IVa. In summary, this study has the best classification accuracy in BCI IV 2a.

In summary, this model can realize the extraction and efficient fusion of spatiotemporal and frequency domain features and obtain the high-precision MI-EEG classification.

C. Ablation studies

We discussed the influence of the multi-scale time-conv fusion unit on classification accuracy in this study. We evaluated multi-scale units individually and compared the result, and it is shown in Table 8. For the multi-scale time-conv fusion unit, the 1st-scale has the highest average accuracy, and the 3rd-scale has the lowest average accuracy. This conclusion showed that the trend of results at the different scales is consistent in the two datasets. The average accuracy by the 1st and 3rd scales was 76.70% and 74.34%, respectively, on BCI IV 2a. The average accuracy by these two different scales was 89.64% and 87.05%, respectively, on HGD. This unit can extract the multi-scale temporal and spatial EEG features, and these features are useful for the classification of MI-EEG. The cascades of these three scales obtained the best results.

We discussed the influence of the frequency-conv fusion unit with the different frequency bands on the classification accuracy. We divided MI-EEG signal into five frequency bands, which are 0-4Hz (δ wave), 4-8Hz (θ wave), 8-12Hz (α wave), 13-30Hz (β wave), and 30-50Hz (γ wave). We evaluated the different frequency bands and compared the classification results, and it is shown in Table 9. The average accuracy of the α band is 78.93% and 90.63% on BCI IV 2a and HGD, respectively. α band

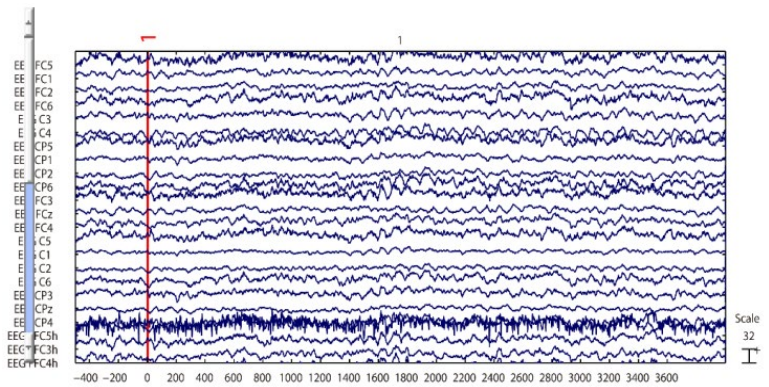
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

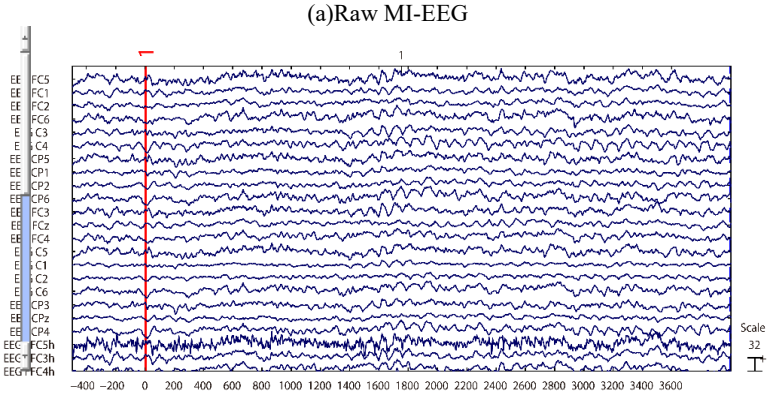
has the highest average accuracy of these five different bands. This result proved that the proposed model has the best classification result in this alpha wave band. The frequency bands represent the different frequency characteristics of MI-EEG, which are useful for classifying MI-EEG, and the feature fusion of the different frequency bands produces the best result. We can conclude that the fusion of these five frequency band features is better than the single frequency band feature.

D. Feature visualization

The visualization of MI-EEG signal features is the basis for our analysis of whether MI-EEG is divisible. The time-frequency diagrams by Daubechies order-4 (Db4) wavelet are shown in Fig. 5. Fig. 5(a) is the time-domain signal of the MI-EEG C3-C4 channel, and Fig. 5(b) is the frequency diagram of the C3-C4 channel, and it shows that there are significant differences in the frequency domain. We used t-SNE to draw the feature distribution for several different models in the ninth subject training set on the two public datasets, and it is shown in Fig.6. The up half is the feature distribution results on the BCI IV 2a, and the button half is the feature distribution results on the HGD, and it contains four classes including left(0), right(1), feet(2), and tongue(3), respectively. This method works by clustering features of the same class together. The results of the feature distributions are shown in Fig. 6, and we can find that the feature visualizations from the EEGNET, DeepConvNet, and ShallowConvNet are relatively ambiguous in contrast to our MSFNet, since the MSFNet model can extract both temporal and spectral features from EEG signals. We can conclude that the four types of features within the two datasets are separable across models and have good separability. It can also be found that different models have different levels of feature aggregation and that MSFNet can utilize time-frequency feature fusion for the MI-EEG classification.



(a)Raw MI-EEG



(b)Preprocessing Results by MSPCA

Fig.4. Preprocessing results

Table 3. Classification accuracy using 5-fold cross validation.

Fold	Training Acc	Testing Acc	F1	K
1	90.24%	83.78%	0.834	0.791
2	91.02%	83.43%	0.836	0.793
3	91.06%	83.24%	0.835	0.794
4	90.01%	84.68%	0.830	0.790
5	89.24%	83.38%	0.823	0.789
Avg.	90.31%	83.68%	0.832	0.791

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

Table 4. Comparison of the results of our method with the other three methods in the BCI IV 2a.

Subject	EEGNet			DeepConvNet			ShallowConvNet			**MSFNet**		
	Acc	F1	K	Acc	F1	K	Acc	F1	K	**Acc**	**F1**	**K**
A01	80.56%	0.803	0.757	68.75%	0.674	0.623	66.32%	0.663	0.596	83.68%	0.832	0.791
A02	54.51%	0.547	0.473	47.57%	0.474	0.405	48.61%	0.458	0.415	74.15%	0.721	0.710
A03	90.28%	0.903	0.874	80.21%	0.800	0.752	83.33%	0.831	0.789	90.68%	0.905	0.876
A04	62.50%	0.624	0.556	72.57%	0.724	0.665	51.74%	0.516	0.446	76.90%	0.746	0.675
A05	73.96%	0.739	0.681	67.71%	0.671	0.611	64.58%	0.637	0.578	78.83%	0.754	0.684
A06	56.25%	0.564	0.491	43.40%	0.425	0.365	47.22%	0.469	0.402	68.04%	0.631	0.614
A07	72.57%	0.720	0.665	71.18%	0.709	0.649	73.61%	0.727	0.677	88.30%	0.851	0.848
A08	77.08%	0.773	0.716	72.22%	0.722	0.661	78.13%	0.777	0.728	79.52%	0.774	0.751
A09	82.29%	0.817	0.777	64.93%	0.642	0.581	69.44%	0.699	0.630	84.16%	0.835	0.738
Avg.	72.22%	0.721	0.666	65.39%	0.649	0.590	64.78%	0.642	0.585	**80.47%**	**0.783**	**0.743**

Table 5. Comparison of the results of our method with the other three methods in the HGD.

Subject	EEGNet			DeepConvNet			ShallowConvNet			**MSFNet**		
	Acc	F1	K	Acc	F1	K	Acc	F1	K	**Acc**	**F1**	**K**
1	91.25%	0.912	0.887	85.62%	0.840	0.794	90.63%	0.906	0.879	95.00%	0.950	0.934
2	80.00%	0.795	0.750	91.87%	0.919	0.895	90.00%	0.899	0.871	92.00%	0.921	0.900
3	97.50%	0.975	0.967	98.75%	0.987	0.983	96.88%	0.969	0.959	96.88%	0.969	0.959
4	85.62%	0.850	0.817	90.63%	0.906	0.879	96.25%	0.962	0.951	96.25%	0.962	0.952
5	94.37%	0.944	0.926	95.63%	0.956	0.943	91.87%	0.920	0.895	96.88%	0.969	0.959
6	89.38%	0.894	0.863	86.25%	0.862	0.825	89.38%	0.893	0.863	91.87%	0.919	0.895
7	90.63%	0.906	0.879	91.87%	0.919	0.895	88.12%	0.881	0.848	93.13%	0.931	0.910
8	93.13%	0.931	0.910	94.37%	0.944	0.926	94.37%	0.943	0.926	95.63%	0.956	0.943
9	94.37%	0.943	0.926	93.75%	0.937	0.918	96.88%	0.968	0.959	95.00%	0.950	0.934
10	90.00%	0.899	0.871	89.38%	0.893	0.863	88.75%	0.887	0.855	91.38%	0.904	0.873
11	80.00%	0.803	0.750	78.75%	0.792	0.735	75.62%	0.755	0.699	85.13%	0.853	0.817
12	91.25%	0.911	0.887	94.37%	0.944	0.926	94.37%	0.944	0.926	95.25%	0.952	0.941
13	93.75%	0.937	0.918	93.75%	0.937	0.918	91.87%	0.919	0.895	93.75%	0.937	0.918
14	76.25%	0.762	0.707	84.38%	0.842	0.802	90.00%	0.899	0.871	91.63%	0.896	0.877
Avg.	89.11%	0.890	0.861	90.67%	0.906	0.879	91.07%	0.910	0.886	**93.56%**	**0.933**	**0.915**

Accuracy retains four significant digits, and F1 score and kappa value retains three valid digits.

Table 6. Number of model parameters.

	Method	Number of parameters(thousand)	Training Time(h)
BCI IV 2a	EEGNet	3.78	2.8
	DeepConvNet	200	7.1
	ShallowConvNet	60.7	4.6
	MSFNet	**13.3**	**3.5**
HGD	EEGNet	4.13	4.5
	DeeConvNet	214	8.2
	ShallowConvNet	95.9	7.4
	MSFNet	**14.7**	**5.7**

Table 7. The classification results of the state-of-the-art methods.

	Method	Year	Acc(%)	K
BCI IV 2a	C2CM[20]	2018	74.46	0.659
	CP-MixedNet[23]	2019	73.20	0.643
	EA-CSP-LDA[31]	2020	73.53	0.647
	TS-SEFFNet[38]	2021	74.71	0.663
	FBMSNet[39]	2023	79.17	0.764
	FBSCN [40]	2023	78.20	0.758
	MSFNet	**2023**	**80.47**	**0.783**
HGD	CP-MixedNet[23]	2019	93.00	0.907
	Residual Net[20]	2019	88.90	0.852
	FBCSP[30]	2019	90.90	0.879
	TS-SEFFNet[38]	2021	93.25	0.910
	MSFNet	**2023**	**93.56**	**0.933**

Table 8. Influence on multi-scale time-conv fusion unit on classification accuracy

	Method	Average Acc(%)
BCI IV 2a	1st-scale	76.70
	2nd-scale	76.58
	3rd-scale	74.34
	MSFNet	**80.47**
HGD	1st-scale	89.64
	2nd-scale	89.42
	3rd-scale	87.05
	MSFNet	**93.42**

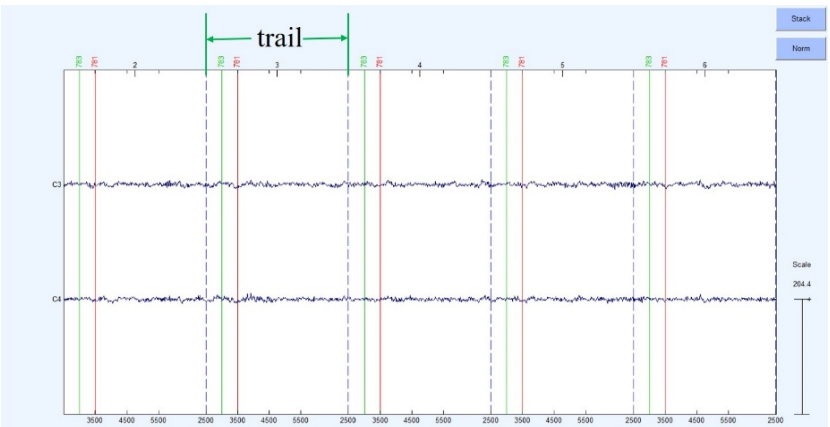
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

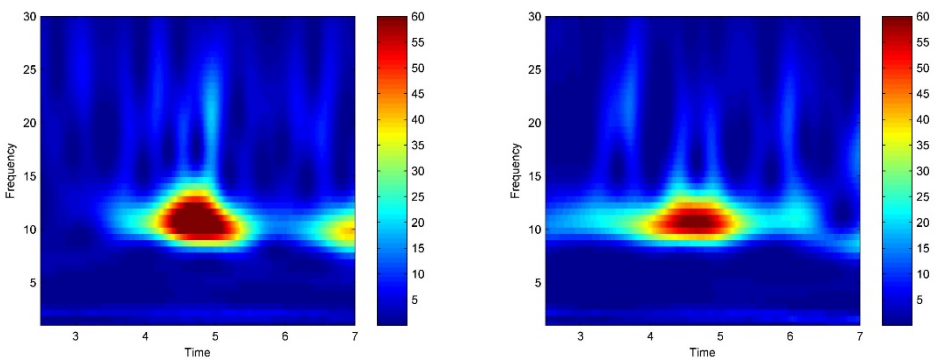
Table 9. Influence on multi-scale frequency-conv fusion unit on classification accuracy

	Subject	δ	θ	α	β	γ	**Our method**
BCI IV 2a	A01	81.60	83.68	81.25	81.94	82.99	**83.68**
	A02	74.93	77.01	78.40	75.97	73.19	**74.15**
	A03	88.54	87.85	88.89	89.58	88.19	**90.68**
	A04	68.40	70.49	70.83	68.68	68.40	**76.90**
	A05	70.49	70.14	72.92	73.96	73.61	**78.83**
	A06	61.81	60.07	66.32	64.58	61.46	**68.04**
	A07	88.19	85.42	86.11	85.76	86.46	**88.30**
	A08	78.82	77.08	78.13	78.47	77.43	**79.52**
	A09	85.07	80.21	79.51	83.33	80.56	**84.16**
	Avg.	78.43	77.77	78.93	78.92	77.81	**80.47**
HGD	1	91.87	91.87	92.50	91.87	91.25	**95.00**
	2	87.50	91.87	88.75	89.38	91.87	**92.00**
	3	96.25	95.00	97.50	95.00	94.37	**96.88**
	4	92.50	91.25	91.87	88.12	88.75	**96.25**
	5	96.25	95.63	96.25	96.88	96.25	**96.88**
	6	87.50	88.12	89.38	88.75	90.00	**91.87**
	7	89.38	88.12	89.38	90.00	88.75	**93.13**
	8	93.13	93.13	94.37	92.50	95.00	**95.63**
	9	94.37	94.37	94.37	94.37	95.63	**95.00**
	10	86.25	86.88	88.75	84.38	88.12	**91.38**
	11	74.38	77.50	77.50	78.13	76.25	**85.13**
	12	91.25	90.00	89.38	91.25	88.12	**95.25**
	13	92.50	94.37	92.50	93.13	91.87	**93.75**
	14	79.37	83.13	86.25	82.50	82.50	**91.63**
	Avg.	89.46	90.09	90.63	89.73	89.91	**93.56**

The result for each frequency band is the result of removing the other four bands from the overall network structure. All results in the table are percentages.



(a) Time-domain signal MI-EEG



(b) Time-frequency chart of MI-EEG C3 channel (c) Time-frequency chart of MI- EEG C3 channel

Fig.5. MI-EEG time and frequency-domain features

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

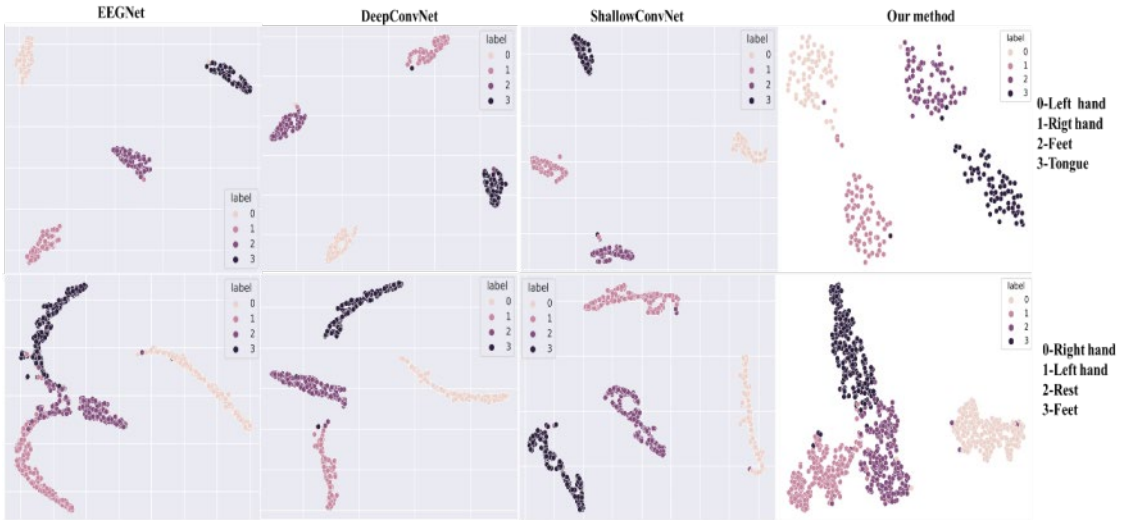


Fig.6. Feature distribution of the ninth subject for the two datasets using the t-SNE method

IV. Discussion

In this study, we adopted EEGNet, DeepConvNet, and ShallowConvNet as the single network models to demonstrate the effectiveness of MSFNet model. The average accuracy, kappa value, and F1 score were the highest on the BCI IV 2a and HGD. The classification results of EEGNet, DeepConvNet, and ShallowConvNet were consistent with the literature [21]. The multi-scale time-space frequency fusion features are more advantageous in classifying MI-EEG than the time-space features. We compared the number of the different model parameters and demonstrated that MSFNet is simpler than DeepConvNet and ShallowConvNet.

We compared the classification results of MSFNet with the state-of-the-art methods [20], [23], [30-31], [38-40]. The average accuracy and kappa value of MSFNet were the highest on BCI IV 2a and HGD. MSFNet and literature [38] used time-frequency features fusion to classify MI-EEG, and the average accuracy and kappa value are higher than the other literature. There are several similar studies on MI-EEG classification. Literature [16] proposed the MI-EEG classification method and tested the validity of the proposed model on the large subject BCI GigaDB dataset, which comprises of 52 normal subjects. Graphical features are one of the newest approaches for identifying underlying patterns of EEG signals, and literature [14] used the Poincaré plot of EEG signals and graphical features in the DWT domain to complete the depression EEG classification. In summary, this study has the best classification accuracy in BCI IV 2a.

We completed the ablation experiment by analyzing the effects of the different time-space and frequency bands. We discussed the influence of multi-scale time-conv fusion unit and demonstrated that the cascades of the three scales produced the best results. We discussed the influence of the multi-scale frequency-conv fusion unit and demonstrated that the five frequency features fusion achieved better results than the individual frequency bands. The conclusion also demonstrated the strong robustness of the proposed model.

In this study, we realized the visualization of the features. We used t-SNE to draw the feature distribution for several different models in the ninth subject training set on the two public datasets by clustering features of the same class together. The conclusion demonstrated the separability of the multi-scale time-space frequency fusion features, and fusion features were more advantageous in the MI-EEG classification.

A small number of the public MI-EEG database was used in this study, and BCI IV 2a and HGD datasets were tested, which is the limitation of this study. Testing more samples on more public datasets will be the future of this study. The new deep learning models and multi-modal features fusion are the possible solutions for the MI-EEG classification, and the graphical features may be another way to realize the efficient classification of MI-EEG.

V. Conclusion

In this study, we proposed a multi-scale space-time frequency fusion network (MSFNet) for MI-EEG classification. We tested and evaluated the MSFNet model using average accuracy, kappa value, and F1 score. We achieved an average accuracy of 80.47%, kappa value of 0.783, and F1 score of 0.743 in BCI IV 2a dataset, and we achieved an average accuracy of 93.56%, kappa value of 0.933, and F1 score of 0.915 in HGD dataset. We compared the classification results of the state-of-the-art methods and demonstrated the superiority of this model. We completed the ablation experiment by analyzing the effects of different scales and frequency bands. In summary, this model can complete MI-EEG classification efficiently.

VI. Data availability

BCI IV 2a and High Gamma Dataset (HGD) are the public MI-EEG datasets in this study. BCI IV 2a dataset

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

(https://www.bbci.de/competition/iv/#download) was provided by the Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology, (Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, Gert Pfurtscheller). HGD dataset (https://github.com/robintibor/high-gamma-dataset) was provided by Intracranial EEG and Brain Imaging lab, Epilepsy Center, Medical Center, University of Freiburg.

VII. Declaration of competing interest

The authors declare that they have no conflict of interest.
Chang Wang: Methodology, Writing-original draft. Yang Wu: Methodology, Software and Writing-original draft. Chen Wang: Resources. Yaning Ren: Visualization. Jiefen Shen: Investigation. Ting Pang: Validation. Chee Seng Chan: Project administration. Wenjie Ren: Project administration. Yi Yu: Funding Acquisition, Resources, Supervision.

VIII. References

[1]X. Zhang, L. Yao, X. Wang, J. Monaghan, D. Mcalpine, Y. Zhang, "A Survey on Deep Learning-based Non-Invasive Brain Signals: Recent Advances and New Frontiers", *IOP Publishing*, 2021.

[2] Devi, K. Renuga, and H. H. Inbarani, "Neighborhood based decision theoretic rough set under dynamic granulation for BCI motor imagery classification", *Journal on Multimodal User Interfaces*, vol. 4, no. 15, pp. 301-21, 2021.

[3] J. Hu, D. Xiao, and Z. Mu, "Application of Energy Entropy in Motor Imagery EEG Classification", *International Journal of Digital Content Technology & Its Applications*, vol. 3, no. 2, pp. 83-90, 2009.

[4] C.C. Chang, C.J. Lin, "LIBSVM: A library for support vector machines", *ACM Transactions on Intelligent Systems and Technology,* no. 2, pp. 1-27, 2011.

[5]C. Vidaurre, M. Kawanabe, P. von Bünau, B. Blankertz, K.R. Müller, "Toward Unsupervised Adaptation of LDA for Brain–Computer Interfaces", *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 587-97, 2011.

[6] M. Li, J. Han, J. Yang, "Automatic feature extraction and fusion recognition of motor imagery EEG using multilevel multi-scale CNN", *Medical & Biological Engineering & Computing*, vol. 59, no. 10, pp. 2037-50, 2021.

[7] K.K. Keerthi, K.P. Soman, "CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image", *Biomedical Engineering Letters*, no. 11, pp. 235-47, 2021.

[8] B. Blankertz, R. Tomioka, S. Lemm, "Kawanabe M, Muller K. Optimizing Spatial filters for Robust EEG Single-Trial Analysis", *IEEE Signal Processing Magazine*, no. 25, pp. 41-56, 2008.

[9] A.K. Keng, C.Z. Yang, W. Chuanchu, et al, "Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b", *Frontiers in Neuroscience*, vol. 6, no. 39, pp. 1-9, 2012.

[10]H. Higashi, T. Tanaka, "Simultaneous Design of FIR Filter Banks and Spatial Patterns for EEG Signal Classification", *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 1100-1110, 2013.

[11]G. Dornhege, "Combined optimization of spatial and temporal filters for improving brain-computer interfacing", *IEEE transactions on bio-medical engineering,* vol. 53, no. 11, pp. 2274-2281, 2006.

[12] N.S. Malan, S. Sharma, "Motor Imagery EEG Spectral-Spatial Feature Optimization Using Dual-Tree Complex Wavelet and Neighbourhood Component Analysis", *IRBM*, vol. 43, no. 3, pp. 198-209, 2022.

[13] M.T. Sadiq, H. Akbari, S. Siuly, et al, "Alcoholic EEG signals recognition based on phase space dynamic and geometrical features", *Chaos Solitons Fractals*, vol. 158, 2022.

[14] H. Akbari, M.T. Sadiq, M. Payan, et al, "Depression Detection Based on Geometrical Features Extracted from SODP Shape of EEG Signals and Binary PSO", *Traitement du Signal*, vol. 38, no. 1, pp. 13-26, 2021.

[15] X. Yu, M.Z. Aziz, M.T. Sadiq, et al, "A New Framework for Automatic Detection of Motor and Mental Imagery EEG Signals for Robust BCI Systems", *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2021.

[16] M.T. Sadiq, X. Yu, Z. Yuan, et al, "Towards the Development of Versatile Brain-Computer Interfaces", *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 4, pp. 1-15, 2021.

[17]L. Yang, Y. Song, K. Ma, L. Xie, "Motor Imagery EEG Decoding Method Based on a Discriminative Feature Learning Strategy", *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 29, pp. 368-379, 2021.

[18]A. Sors, S. Bonnet, S. Mirek, L. Vercueil, J.F. Payen, "A convolutional neural network for sleep stage scoring from raw single-channel EEG", *Biomedical Signal Processing and Control*, vol. 42, no. 107-114, 2018.

[19]T. Tezuka, D. Kumar, S. Singh, I. Koyanagi, T. Naoi, M. Sakaguchi, "Real-time, automatic, open-source sleep stage classification system using single EEG for mice", *Scientific Reports*, vol. 11, no. 1, pp. 1-8, 2021.

[20]R.T. Schirrmeister, J.T. Springenberg, L.D.J. Fiederer, et al, "Deep learning with convolutional neural networks for EEG decoding and visualization", *Human Brain Mapping*, vol. 38, no. 11, pp. 5391-5420, 2017.

[21] X. Ma, S. Qiu, C. Du, J. Xing, H. He, "Improving EEG-Based Motor Imagery Classification via Spatial and Temporal Recurrent Neural Networks", *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE*

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204

Engineering in Medicine and Biology Society. Annual Conference, 2018.

[22]S. Sakhavi, C. Guan, S. Yan, "Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5619-5629, 2018.

[23]Y. Li, X.R. Zhang, B. Zhang, M.Y. Lei, W.G. Cui, Y.Z. Guo, "A Channel-Projection Mixed-Scale Convolutional Neural Network for Motor Imagery EEG Decoding", *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 27, no. 6, pp. 1170-1180, 2019.

[24]H. He, D. Wu, "Transfer Learning for Brain–Computer Interfaces: A Euclidean Space Data Alignment Approach", *IEEE transactions on bio-medical engineering*, vol. 67, no. 2, pp. 399-410, 2020.

[25]B. Sun, X. Zhao, H. Zhang, R. Bai, T. Li, "EEG Motor Imagery Classification With Sparse Spectrotemporal Decomposition and Deep Learning", *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 541-551, 2021.

[26]K. Zhang, N. Robinson, S.W. Lee, C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network", *Neural Networks: The Official Journal of the International Neural Network Society*, vol. 136, pp. 1-10, 2021.

[27] J.S. Huang, W.S. Liu, B.Yao, Z.X. Wang, S.F. Chen, W.F. Sun, "Electroencephalogram-Based Motor Imagery Classification Using Deep Residual Convolutional Networks", *Frontiers in Neuroscience*, vol. 15, pp. 1-8, 2021.

[28]A. Kumar, C. Chakraborty, B. Gupta, "Secure Transmission of EEG Data Using Watermarking Algorithm for the Detection of Epileptical Seizures", *Traitement du Signal: signal image parole*, vol. 2, no. 38, pp. 473-479, 2021.

[29] A. Kumar, K. Abhishek, C. Chakraborty, et al, "Deep Learning and Internet of Things Based Lung Ailment Recognition Through Coughing Spectrograms", *IEEE Access*, vol. 9, pp. 95938-95948, 2021.

[30]Y. Li, L. Guo, Y. Liu, J. Liu, F. Meng, "A Temporal-Spectral-Based Squeeze-and- Excitation Feature Fusion Network for Motor Imagery EEG Decoding", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1534-45, 2021.

[31] K. Liu, M. Yang, Z. Yu, et al, "FBMSNet: A Filter-Bank Multi-Scale Convolutional Neural Network for EEG-Based Motor Imagery Decoding", *IEEE transactions on bio-medical engineering*, vol. 70, no. 2, pp. 436-445, 2022.

[32] J. Chen, D. Wang, W. Yi, et al, "Filter bank sinc-convolutional network with channel self-attention for high performance motor imagery decoding", *Journal of Neural Engineering*, vol. 20, no. 2, pp. 026001, 2023.

[33]M.T. Sadiq, X.J. Yu, Z.H. Yuan, et al, "Motor Imagery BCI Classification Based on Multivariate Variational Mode Decomposition", *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 5, pp. 1177-1189, 2022.

[34] M.T. Sadiq, X. Yu, Z. Yuan, et al, "Motor imagery BCI classification based on novel two-dimensional modelling in empirical wavelet transform", *Electronics Letters*, pp. 1367-1369, 2020.

[35] M.T. Sadiq, X. Yu, Z. Yuan, et al, "Motor Imagery EEG Signals Decoding By Multivariate Empirical Wavelet Transform Based Framework For Robust Brain-Computer Interfaces", *IEEE Access*, vol. 7, pp. 171431-171451, 2019.

[36]D. Freer, G.Z. Yang, "Data augmentation for self-paced motor imagery classification with C-LSTM", *Journal of Neural Engineering*, vol. 17, no. 1, pp. 1-15, 2020.

[37] Z. Deng, P. Xu, L. Xie, et al, "Transductive Joint-Knowledge-Transfer TSK FS for Recognition of Epileptic EEG Signals", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1481-1494, 2018.

[38]Y. Zhang, B. Liu, X. Ji, D. Huang, "Classification of EEG Signals Based on Autoregressive Model and Wavelet Packet Decomposition", *Neural Processing Letters*, no. 45, pp. 365–78, 2017.

[39]Y. Li, Y. Liu, W.G. Cui, Y.Z. Guo, H. Huang, Z.Y. Hu, "Epileptic Seizure Detection in EEG Signals Using a Unified Temporal-Spectral Squeeze-and-Excitation Network", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 99, pp. 782-94, 2020.

[40]M.C. Cieslak, A.M. Castelfranco, V. Roncalli, P.H. Lenz, D.K. Hartline, "t-Distributed Stochastic Neighbor Embedding (t-SNE): A tool for eco-physiological transcriptomic analysis", *Mar Genomics*, no. 51, pp. 1-12, 2020.

[41] M.T. Sadiq, X. Yu, Z. Yuan, et al, "Motor Imagery EEG Signals Decoding By Multivariate Empirical Wavelet Transform Based Framework For Robust Brain-Computer Interfaces", *IEEE Access*, no. 99, pp. 1-22, 2019.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3351204



Wang Chang received a Bachelor's degree in Biomedical Engineering from Xinxiang medical university, China, and the Master degree in Biomedical Engineering from University of Science and Technology of China. He is currently associate professor with the School of Medical Engineering, Xinxiang Medical University. His research interests include medical image analysis, brain-computer interface, AI (machine learning and deep learning), and computer aided diagnosis.



Wu Yang received Bachelor's degree in Biomedical Engineering (Medical Device Engineering) and a Master's degree in Biology and Medicine from Xinxiang Medical University. His doctoral degree is currently studying at the School of Electronic and Information Engineering at Tiangong University. His research interests include physiological signal analysis, brain-computer interface and AI (machine learning and deep learning).



Wang Chen received a bachelor's degree in medical imaging technology from Xinxiang Medical University. She is currently studying for a master's degree at Xinxiang Medical University. Her research interests include resting state cerebral functional magnetic resonance imaging, artificial intelligence (machine learning and deep learning), and computer-aided diagnosis.



Ren Yaning received a bachelor's degree in medical imaging technology from Xinxiang Medical university. she is now a graduate student at Xinxiang Medical university. She is currently researching medical image analysis, artificial intelligence (machine learning and deep learning) and computer-aided diagnostics.



Shen Jiefen received a Master degree from Zhengzhou university, China. He mainly engaged in physical electronics and electronic technology in the field of basic and applied research.



Pang Ting received a Bachelor's degree in Computer Science and Technology, Anyang Normal University, China, and the Master degree in Management Science and Engineering, Beijing University of Technology, China, and the Doctoral degree in Computer Science and Information Technology, Universiti Malaya, Malaysia. She is currently Lecture with the School of Medical Engineering, Xinxiang Medical University. Her research interests include Medical Image Analysis and Deep Learning.



Chee Seng Chan is currently a Full Professor with the Faculty of Computer Science and Information Technology, University Malaya, Malaysia. His research interests include computer vision and machine learning with focus on scene understanding. Prof. Chan was the Founding Chair for IEEE Computational Intelligence Society (Malaysia Chapter), Organising Chair for the Asian Conference on Pattern Recognition (2015), General Chair for the IEEE Workshop on Multimedia Signal Processing (2019), and IEEE Visual Communications and Image Processing (2013). He is a Chartered Engineer registered under the Engineering Council, U.K. He was the recipient of several notable awards, such as Young Scientist Network-Academy of Sciences Malaysia in 2015 and the Hitachi Research Fellowship in 2013.



Ren Wenjie received a Ph.D. degree from Zhengzhou university, China. He is doctor of hospital administration and dean of Xinxiang Medical University. His research interests big data, and computer aided diagnosis.



Yu Yi received a Master and Ph.D. degree in Biomedical Engineering from Zhejiang university, China. He is currently professor and Director with the School of Medical Engineering, Xinxiang Medical University. He is the Director with Henan International Joint Laboratory of Neural information analysis and Drug intelligent design. His research interests brain-computer interface, and computer aided diagnosis.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4