IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 January 2024, accepted 26 February 2024, date of publication 29 February 2024, date of current version 6 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3371904

RESEARCH ARTICLE

A Study of Unilateral Upper Limb Fine Motor Imagery Decoding Using Frequency-Band Attention Network

TIANYU SHI[1], XIANG GU[2], HUI BI[2], JIDONG LV[1], YAN LIU[3,4], YAKANG DAI[3,4], AND LING ZOU[1,5]

[1]School of Microelectronics and Control Engineering, Changzhou University, Changzhou, Jiangsu 213164, China
[2]School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, Jiangsu 213164, China
[3]Department of Medical Image, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou 215163, China
[4]Suzhou Guoke Kangcheng Medical Technique Company Ltd., Suzhou 215163, China
[5]Key Laboratory of Brain Machine Collaborative Intelligence Foundation of Zhejiang Province, Hangzhou, Zhejiang 310018, China

Corresponding authors: Ling Zou (zouling@cczu.edu.cn) and Yakang Dai (daiyk@sibet.ac.cn)

This work was supported in part by the Project of Jiangsu Key Research and Development Plan under Grant BE2021012-2 and Grant BE2021012-5, in part by Changzhou Science and Technology Bureau Plan under Grant CE20225034, in part by the Key Laboratory of Brain Machine Collaborative Intelligence Foundation of Zhejiang Province under Grant 2020E10010-04, in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX22_3058, and in part by the Human–Machine Intelligence and Interaction International Joint Laboratory Project.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Changzhou University.

ABSTRACT Brain-computer interface (BCI) based motor imagery (MI) can assist stroke patients in upper limb rehabilitation and help restore motor function to a certain extent. However, the classical MI paradigm distinguishes different limbs and cannot effectively meet the needs of upper limb rehabilitation training for patients. Therefore, this paper designed a new paradigm for three motor imagery actions targeting different joints of the unilateral upper limb, and electroencephalogram (EEG) data from 20 healthy participants were collected for research analysis. A deep neural network model combining an attention mechanism for multiple frequency bands and a deep convolutional network were proposed to adaptively assign weight to the EEG data in different frequency bands. Then feature extraction was performed for each frequency band to learn further and to classify features. This model can obtain an average accuracy of 69.2% for the subject-independent case with the triple classification in the designed fine motor imagery (FMI) dataset, which is better than other controlled methods. Furthermore, ablation experiments were conducted for each module, demonstrating the effectiveness of each module. These results manifest the feasibility of our proposed method and the potential of FMI paradigm for BCI, providing a new training tool for upper limb rehabilitation after stroke.

INDEX TERMS Brain–computer interface, upper limb rehabilitation, EEG, fine motor imagery, deep learning.

I. INTRODUCTION

Brain-computer interface (BCI) is a system that provides a pathway between the brain and an external device by decoding brain electrical signals. Electroencephalogram (EEG) - based Motor Imagery (MI) Brain-Computer Interfaces (BCIs) have been extensively studied [1]. In contrast to other control paradigms such as P300 and SSVEP, motor imagery-based BCIs enable control of external devices through spontaneous brain activity without the need for additional stimuli [2], thereby expressing the user's intrinsic motor intention.

Stroke is one of the leading causes of mortality and disability worldwide [3], with survivors often experiencing motor impairments and various forms of disability in their daily

The associate editor coordinating the review of this manuscript and approving it for publication was Qichun Zhang.

lives [4]. Hemiparesis, resulting from stroke, is a significant disability, with more than 85% of patients suffering residual upper limb impairments, greatly impacting their quality of life [5]. Research suggests that Motor Imagery-Based Brain-Computer Interface (MI-BCI) interventions offer a promising avenue for stroke patients' rehabilitation [6].

Classical motor imagery paradigms are typically employed to distinguish among four body parts: tongue, left hand, right hand, and foot [7], [8]. However, this mismatch between the imagined actions and the targeted rehabilitation areas fails to meet the specific needs of patients. In recent years, researchers have explored the decoding of fine motor imagery (FMI) of different segments of the unilateral upper limb. Ofner et al. [9] devised a paradigm involving six different actions (elbow flexion, elbow extension, forearm pronation, forearm supination, hand closing, and hand opening), with each motor imagery lasting for 3 seconds and being performed 60 times by 15 healthy subjects. Suwannarat et al. [10] designed three sets of hand movements (hand opening and closing, forearm rotation, wrist flexion and extension) with each motor imagery lasting 4 seconds, and engaged 11 healthy participants in 12 weeks. Ma et al. [11] introduced four distinct actions (hand, wrist, elbow, and shoulder), with each motor imagery session lasting 4 seconds and being conducted 100 times by 20 healthy subjects for the study.

For MI-BCI systems, one of the significant challenges lies in the choice of decoders. Since Ramoser et al. [12] introduced Common Spatial Patterns (CSP), various derived methods, combined with Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA), have been widely utilized in MI decoding due to their high discriminative power and low computational complexity [13]. In recent years, with the rapid advancement of deep learning, an increasing number of researchers have incorporated deep learning algorithms into MI decoding. Deep neural networks can learn to leverage the entire dataset, enabling the development of robust and automated classification systems, supporting end-to-end classification decoding models [14]. Khademi et al. [15] combined Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) and achieved the highest accuracy of 92% on the BCI Competition IV-2a dataset. Amin et al. [16] introduced MCNN, which fuses different features and architectures, and achieved an accuracy of 95.4% on the High Gamma dataset. Roy [17] proposed an efficient multiscale CNN (MS-CNN) with a classification accuracy of 93.74% on the BCI Competition IV-2a dataset. Zhang et al. [18] designed a five-class motor imagery task involving imagining actions for the left hand, right hand, both hands, both feet, and rest. They concurrently collected EEG and fNIRS data. Additionally, they constructed a multimodal MI decoding neural network. This network utilized convolutional networks for feature extraction from both EEG and fNIRS data. Subsequently, an attention mechanism was employed to integrate these two sets of features before the final classification process took place. However, most of these methods were validated on traditional MI datasets and have not been experimentally tested for FMI. However, in the field of FMI, Edelman et al. [19] employed EEG source imaging techniques. They utilized the Colin27 MRI average brain for data domain transformation and achieved remarkable results in classifying four kinds of wrist MI movements. Additionally, in 2019, the team successfully applied the ESI method to online BCI systems for four MI tasks: left hand, right hand, both hands, and rest [20]. This had a significant impact on the development and implementation of non-invasive BCI systems.

To meet stroke patients' upper limb rehabilitation needs, this paper designed movements for different joints of the unilateral upper limb. The FMI paradigm was used instead of the traditional MI paradigm so that the imagined actions correspond to the rehabilitation sites to better fit the actual needs of the hemiplegic patients. On the basis of previous research [21], new FMI actions were added and the classification categories were expanded. Meanwhile, the attention mechanism for multiple frequency bands was combined with deep convolutional networks to decode and to classify the FMI-EEG signals. The experiment results show that the average accuracy is 69.2% in the three-class FMI scenes, which is better than the selected comparison methods.

II. MATERIALS

A. SUBJECTS

We recruited 20 healthy participants (17 males and 3 females) for this experiment, aged between 22 to 26 years and all right-handed. They did not have any MI-based BCI experience before. They were all mentally healthy, with no head trauma, and had normal or corrected vision. Prior to the experiment, a comprehensive overview of the experimental procedures was provided for the participants. After the experiment, participants were interviewed about their experiences to ensure the full completion of the motor imagery experiment. All human experiments were conducted in accordance with the principles of the Helsinki Declaration, and the study received approval from the ethics committee of Changzhou University.

B. EXPERIMENTAL PARADIGM

Prior to the commencement of the experiment, participants were instructed to sit comfortably in a chair with their hands placed naturally on the armrests. They were encouraged to keep their entire body relaxed, maintain focused attention throughout the trial, avoid unnecessary limb movements, and minimize the frequency of blinking during the experiment.

The overall experimental procedure is illustrated in Fig. 1. Prior to the motor imagery experiment, a brief training session was provided to help the participants become familiar with the process. Once it was confirmed with the participants that they were ready to begin the experiment, the formal motor imagery task commenced. One session consisted of

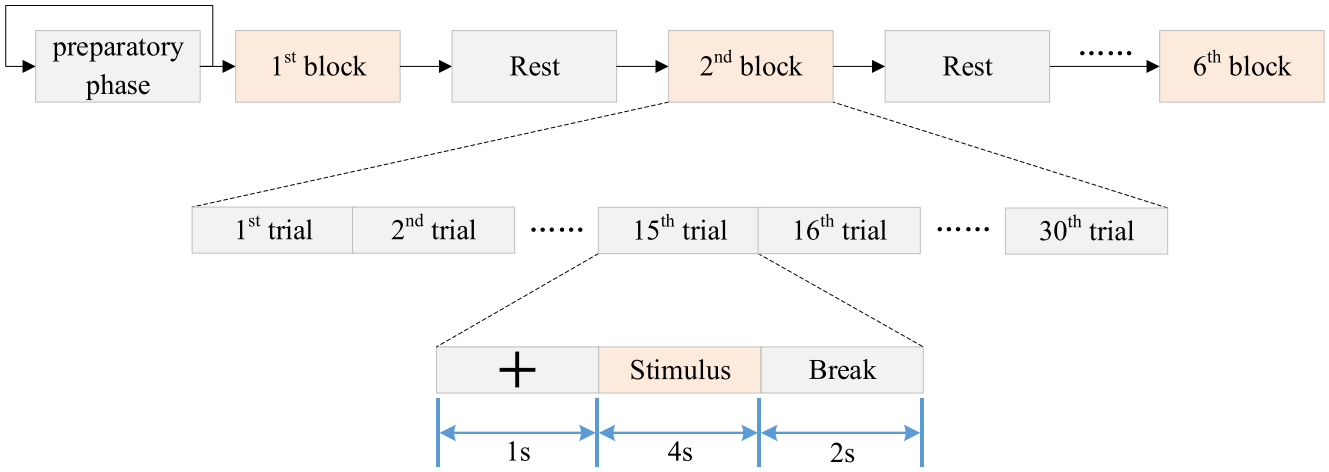


FIGURE 1. Experimental paradigm flow. One session contains six blocks, with a certain amount of rest time between each block. Each block includes 30 trials, and each trial lasts 7 seconds.

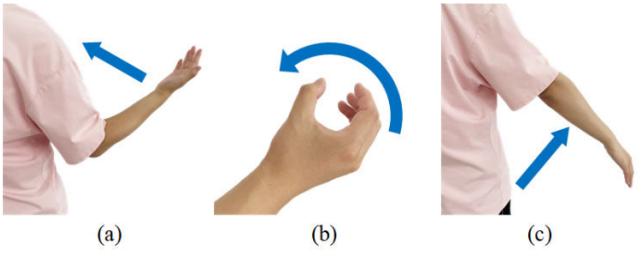


FIGURE 2. Stimulating pictures of fine-motor imagination. Picture (a) elbow flexion, picture (b) fist-clenching, picture (c) shoulder rotation.

6 blocks, with rest intervals in between, and each block comprised 30 trials. At the beginning of each trial, a central fixation cross appeared on the screen for 1 second, signaling the impending motor imagery task. Subsequently, the fixation cross disappeared, and a specific motor imagery cue was presented, as shown in Fig. 2, representing actions such as elbow flexion, fist clenching, and shoulder rotation. Each image was displayed for 4 seconds, followed by a 2-second rest before the next trial began. After each session, participants are queried to ensure sustained attention throughout the entire session and adequate engagement in motor imagery.

Participants were instructed to repetitively engage in motor imagery of the action depicted in the stimulus image during the 4-second duration when the image was presented. Consequently, it can be believed that during these 4 seconds, participants' motor imagery was continuous and roughly similar in time. The order of the image stimuli was randomized, but each image appeared an equal number of times. Each participant undergoes a total of 3 sessions, with each session comprising 6 blocks, and each block consisting of 30 trials. Consequently, each participant engages in 540 trials in the motor imagery experiment. We employed Qt to develop the entire paradigm workflow, utilizing "PreciseTimer" to accurately control the onset and offset times of each stimulus. At the moment of image presentation, we conducted event marking to facilitate subsequent analysis and processing of the EEG signals.

C. DATA ACQUISITION

The data acquisition equipment used in the experiment was the Neuracle NeuSen W series wireless electroencephalography (EEG) signal acquisition system, which featured a total of 59 effective channels. EEG electrodes were placed in accordance with the international 10-20 system. The sampling rate was set at 1000 Hz. Throughout the data collection process, efforts were made to maintain electrode impedances below $10k\Omega$. The EEG experiment was conducted in a quiet and noise-free environment to minimize external interference

D. EEG PRE-PROCESSING

The raw EEG data underwent a basic preprocessing pipeline. Firstly, a 50 Hz notch filter was applied to remove power-line interference. Subsequently, the data was bandpass filtered from 1 to 45 Hz using an eighth-order Butterworth digital filter represented in second-order sections (SOS) format to ensure filter stability. Next, a whole-brain average re-referencing was employed to mitigate the impact of specific reference electrodes. Finally, independent component analysis (ICA) was used and supplemented with manual review to remove EOG, EMG, and ECG and other artifactual components. The data was segmented into -1 to 4 seconds epochs, with baseline correction applied using data from -1 to 0 seconds. According to our self-designed paradigm program, there exists a marker in the EEG dataset precisely at the moment when the image appears. We define this moment as time point 0, and all segmentation operations performed on the data are based on this zero-time reference point.

III. METHODS

The attention-based end-to-end unilateral upper limb fine motor imagery classification network structure is proposed in Fig. 3. The raw EEG data were pre-processed and subdivided into different frequency bands to reconstitute the data set. The multi-band EEG data were input into the frequency-band attention module to allow the network to learn the feature information among different

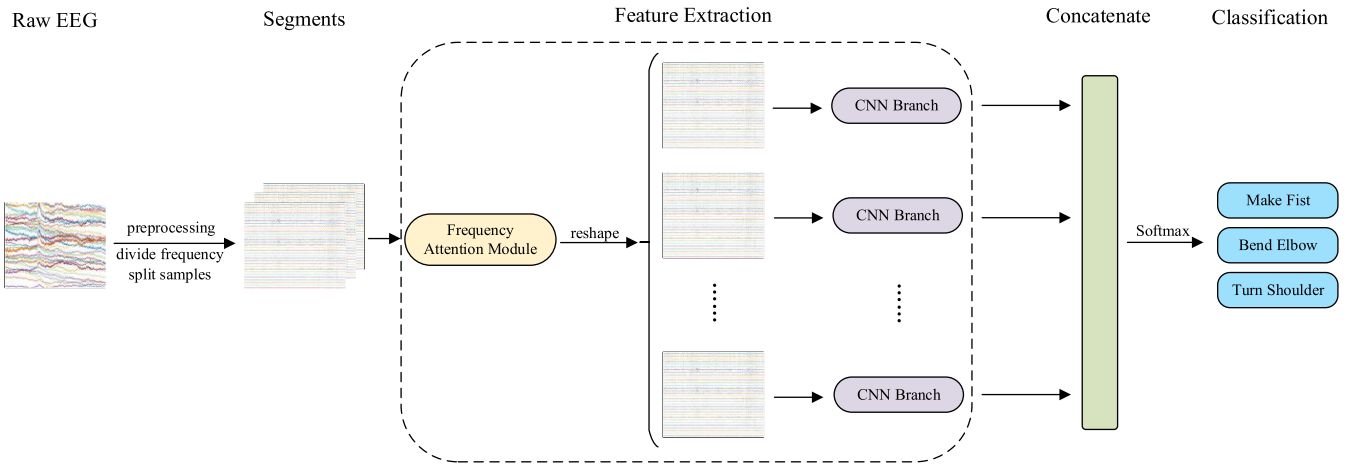


FIGURE 3. Structure of the overall classification network system. The raw EEG data is preprocessed and divided into frequency bands and then input to the frequency band attention module. Then the EEG of each frequency band is split, and each enters and exits the CNN convolutional neural network for further feature extraction. Finally, the features are spliced and classified through the SoftMax layer.

frequency bands. Then, the signals of each frequency band were split and input into the CNN branch for feature extraction of details. Finally, the features were concatenated together and classified by SoftMax to output the final results. All participant data were independently employed, with no amalgamation of data for inter-group training and testing. These steps are further explained in the following subsections.

A. BAND ATTENTION MODULE

1) FREQUENCY BAND SELECTION

The amplitude and frequency values of EEG signals can be used to distinguish a various range of physiological activities. At present, EEG frequencies are mainly determined according to the following frequency bands [7], [14], Delta (1-3 Hz), Theta (4-7 Hz), Alpha (8-13 Hz), Beta (13-30 Hz), Gamma (above 30 Hz). According to many recent studies [8], [22], the brain activities associated with motor imagery usually exist in the Alpha, Beta, and Gamma frequency bands. A few studies [16], [23] extended the frequency band to a broader range. Therefore, in order to study the influence of different frequency bands of EEG signals on the classification results of fine-motor imagery, five frequency bands were selected in this paper, 1-4 Hz, 4-8 Hz, 8-13 Hz, 13-31 Hz, and 31-40 Hz. The signal was subjected to frequency band segmentation using an 8th-order IIR Butterworth filter. With a sampling rate of 1000 Hz, a bandpass filter was employed, and the passbands were chosen based on the aforementioned five frequency ranges. The filter was represented in SOS form to ensure the stability of the filter.

2) ATTENTION MODULE SELECTION

After pre-processing and frequency band extraction, the EEG signal of one trial can be expressed as $E \in \mathbb{R}^{N\times C\times T}$, where N is the number of selected frequency bands, C is the number of EEG cap channels, and T is the number of sampling points. Here, the values of N, C are 5 and 59. The value of T is determined according to the following data enhancement method.

Attention mechanisms have been extensively employed in various EEG signal analyses [24], [25]. They enable models to focus on different parts of the input, intelligently assigning weights to different inputs based on the specific task, demonstrating excellent performance across a variety of tasks [26]. Its essence lies in computing attention coefficients for corresponding Q (Query), K (Key), and V (Value) [27], as shown in (1), thereby assigning weights to different information. In this context, the frequency band attention network structure uses the SE network [28], which focuses on the channel relationship and can extract the data correlation features between frequency bands well. Therefore, the frequency bands are used as channels and input into the SE network.

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \tag{1}$$

The SE network first compresses the global spatial information into channels, which is done by global average pooling. The expression is shown in (2), z_n represents the general information of the nth channel and u_n represents the 2D EEG signal of the nth channel, which can be expressed as $u_n \in \mathbb{R}^{C\times T}$, $(n = 1, \ldots, N)$.

$$z_n = F_{sq}(u_n) = \frac{1}{C \times T}\sum_{i=1}^{C}\sum_{j=1}^{T} u_n(i, j) \tag{2}$$

Then the compressed channel information, through adaptive activation, is obtained as a weight matrix s, which can be used in. δ is the activation function, W_1 and W_2 are fully connected matrices, and $W_1 \in \mathbb{R}^{\frac{N}{r}\times N}$, $W_2 \in \mathbb{R}^{N\times\frac{N}{r}}$, and r are a hyperparameter to control the feature dimensionality reduction.

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2\delta(W_1 z)) \tag{3}$$

The network structure for calculating each frequency band's weight is shown in Table 1.

TABLE 1. Parameters of the attention network for the calculation of frequency band weight.

Layer	Output
Input	(N, C, T)
GlobalPooling2D	$(N, 1, 1)$
Reshape	(N)
Dense	$(N // r)$
ReLU	$(N // r)$
Dense	(N)
Sigmoid	(N)
Reshape	$(N, 1, 1)$

Finally, the original EEG data are multiplied by the weight s to obtain the feature matrix $\tilde{X}_A = [\tilde{x}_{A,1}, \tilde{x}_{A,2}, \ldots, \tilde{x}_{A,N}] \in \mathbb{R}^{N\times C\times T}$, which can be expressed by (4). Where $\tilde{x}_{A,n}$ represents the weighted feature matrix of the nth channel and s_n represents the weight of the nth channel.

$$\tilde{x}_{A,n} = F_{scale}(u_n, s_n) = s_n u_n \tag{4}$$

B. FEATURE EXTRACTION MODULE

To perform further feature extraction for each data band, $\tilde{X}_A$ is dimensionally reshaped and can be expressed as $\tilde{X}_R \in \mathbb{R}^{N\times 1\times C\times T}$. $\tilde{x}_{R,n} \in \mathbb{R}^{1\times C\times T}, (n = 1, \ldots, N)$ denotes the feature data of the nth frequency band.

The EEGNet [29] model is selected as the primary method for feature extraction of each frequency band data. EEGNet is a compact convolutional neural network model that uses a depth-wise separable convolutional structure, which can effectively and quickly extract feature information from EEG signals. Its network structure parameters are shown in Table 2.

After feature extraction, the feature matrix of the nth frequency band can be expressed as $\tilde{x}_{F,n} \in \mathbb{R}^{N_{class}}$, as shown in (5), where $F_{feature}$ represents the feature extraction process. Finally, the features of all frequency bands $\tilde{X}_F = [\tilde{x}_{F,1}, \tilde{x}_{F,2} \ldots, \tilde{x}_{F,N}]$ are concatenated together and input into the classifier as the final features for classification.

$$\tilde{x}_{F,n} = F_{feature}(\tilde{x}_{R,n}) \tag{5}$$

C. CLASSIFICATION MODULE

After the data have gone through all the feature extraction modules, the obtained features are obtained as $\tilde{X}_F \in \mathbb{R}^{N\times N_{class}}$. After all the features are flattened, the final classification result is output after a fully connected layer and a normalized exponential function layer. The network structure is shown in Table 3.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL EQUIPMENT AND RELATED PARAMETERS

This deep learning network was implemented using the Pytorch framework and was trained with a TESLA V100 graphics card. The network was trained using Adam [30] optimizer, cross-entropy loss function. The learning rate was chosen from [0.01, 0.001, 0.0001], and the hyperparameters in the attention network were selected from [2], [3], and [5]. Dataset splitting used five-fold cross-validation.

In order to expand the dataset as much as possible, data augmentation was performed using down sampling, with augmentation factors selected from the range [2], [4], [8]. A data augmentation example is presented in Fig. 4. Assuming the original data consists of 12 sampling points, a 2 times data augmentation was applied using down sampling, resulting in data segments with 6 sampling points each. Although the number of sampling points per individual segment is reduced, the overall data quantity is doubled. After data augmentation, data segments that originally belonged to the same source may become overly similar. If they were to appear separately in both the training and validation sets, it could lead to overfitting. Therefore, this data augmentation operation was performed after the dataset were partitioned.

B. METHOD OF COMPARISON

1) CLASSIC MACHINE LEARNING

CSP [12] and the CSP-based FBCSP [31] for feature extraction and classification use SVM. Due to the fact that CSP-based methods are designed for binary classification tasks, in the context of the three-class problem addressed in this paper, a one-versus-rest (OVR) approach is employed. The training and application of the SVM classifier are both implemented using the Python Scikit-Learn library to ensure the correctness of the method's implementation.

2) DEEP LEARNING

This paper selected the deep learning network models related to MI in recent years for comparative experiments. Perform network was reproduced according to the method described in the paper. To adapt to the data set, the input layer structure of each network was fine-tuned to ensure that the EEG data preprocessing methods and sampling rates used by each method are consistent and that the deep learning network itself is consistent with the description in its paper.

- MCNN [16]: A multi-layer CNN network structure that integrates CNN networks with different features and structures to improve the classification accuracy of MI-EEG data. The selected frequency bands for data input are 7-13Hz, 13-31Hz, 7-31Hz, and 0-40Hz.
- MSCNN [17]: An efficient, multi-scale CNN for feature extraction and classification of multiple non-overlapping standard frequency band. The selected frequency bands for data input are 1-4Hz, 4-8Hz, 8-13Hz, and 13-30Hz.
- TD-Atten [11]: A model that combines multi-band CSP, attentional mechanism, and long and short-term memory network for the same limb. The data input frequency bands range from 4Hz to 38Hz, with frequency band selection performed using a sliding

TABLE 2. Parameter structure of feature extraction network.

Layer	Size	Output	Options
Input		$(1, C, T)$	
Conv2D	(1, 64)	(F_1, C, T)	padding = same
BatchNorm2D		(F_1, C, T)	
DepthwiseConv2D	(C, 1)	$(D * F_1, 1, T)$	padding = valid, depth = D, max norm = 1
BatchNorm2D		$(D * F_1, 1, T)$	
ELU		$(D * F_1, 1, T)$	
AveragePool2D	(1, 4)	$(D * F_1, 1, T // 4)$	
Dropout		$(D * F_1, 1, T // 4)$	$p = 0.25$ or $p = 0.5$
SeparableConv2D	(1, 16)	$(F_2, 1, T // 4)$	padding = same
BatchNorm2D		$(F_2, 1, T // 4)$	
ELU		$(F_2, 1, T // 4)$	
AveragePool2D	(1, 8)	$(F_2, 1, T // 32)$	
Dropout		$(F_2, 1, T // 32)$	$p = 0.25$ or $p = 0.5$
Flatten		$(F_2 * (T // 32))$	
Dense		N_{class}	

TABLE 3. Classification module network structure.

Layer	Output
Input	(N, N_{class})
Flatten	$N * N_{class}$
Dense	N_{class}
SoftMax	N_{class}

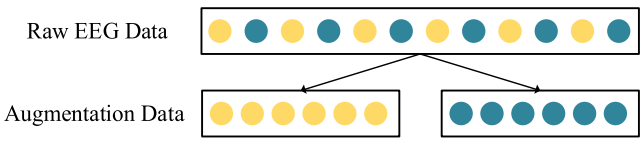


FIGURE 4. A data augmentation example.

window approach. The window size is 4Hz, and the step size is 2Hz, resulting in a total of 16 frequency bands.

- MTFB-CNN [32]: An end-to-end MI-EEG decoding model that integrates attention mechanism and residual network without data pre-processing. The data input frequency band is set to 7-31Hz.
- EEGNet [29]: A compact convolutional neural network model that uses a depth-wise separable convolutional structure, which can effectively and quickly extract feature information from EEG signals. The selected data input frequency band is 8-30Hz.

C. RESULTS

Table 4 shows the classification results of the selected method on the FMI-EEG data of 20 participants. In the table, bold font indicates the best accuracy achieved by each method for a given participant. As shown in Table 4, the method proposed in this study achieved the highest accuracy of 88.89% for participant 11. Since a five-fold cross-validation was employed during the model training process, the average accuracy represents the mean of the highest accuracy achieved in each fold.

Fig. 5 depicts the corresponding box plots. From Fig. 5 and Table 4, it is evident that, in terms of average accuracy on this dataset, the method proposed in this study has shown substantial improvement compared to other methods, achieving the highest improvement of 27.37% and an average accuracy of 69.20%. Furthermore, t-tests were conducted between the proposed method and the baseline CSP+SVM method, with "*" indicating $p < 0.05$ and "**" indicating $p < 0.01$. The results reveal a significant difference in favor of the proposed method compared to the baseline method, highlighting the effectiveness of this approach. We also conducted t-tests between all the methods and the proposed method, and the results are presented in Table 5 The p-values are all less than 0.05, indicating a significant difference between them.

The specific information and comparison of the selected methods are detailed in Table 6. The table lists the year of proposal, average accuracy on this dataset, statistics of trainable model parameters, and evaluation criteria, including kappa and F1 score. As evident from Table 6, the model proposed in this study has fewer trainable parameters compared to other models, enabling faster training and reducing computational resource consumption. Moreover, on this dataset, both kappa and F1 values for the proposed model are significantly higher than those of other methods.

We selected representative participants' FMI-EEG classification results and visualized their confusion matrices,

TABLE 4. Classification accuracies of various methods (in percentage, denoted as * for $p < 0.05$ and ** for $p < 0.01$ in the column of average classification accuracy). The method proposed in this study achieved an average accuracy of 69.20%, representing a remarkable improvement of 27.37% when compared to the selected alternative methods.

	CSP	FBCSP	MCNN	MSCNN	TD-Atten	MTFB-CNN	EEGNet	Proposed
Sub1	46.3	50	48.84	45.14	40.74	50.93	46.53	**55.56**
Sub2	44.44	44.44	54.63	49.77	42.59	49.54	49.31	**56.94**
Sub3	42.59	46.3	43.29	59.03	40.74	54.63	52.66	**62.96**
Sub4	46.3	51.85	59.95	61.34	44.44	51.85	54.05	**74.54**
Sub5	44.44	55.56	57.64	59.72	38.89	61.57	52.66	**62.04**
Sub6	50	51.85	50.93	54.86	44.44	49.07	45.95	**80.79**
Sub7	55.56	62.96	56.48	54.63	38.89	56.48	66.32	**74.07**
Sub8	46.3	46.3	58.56	62.96	40.74	58.8	55.56	**71.06**
Sub9	50	48.15	50.46	50.93	50	50	49.88	**66.9**
Sub10	40.74	42.59	47.69	47.45	40.74	45.37	42.13	**68.75**
Sub11	57.41	61.11	56.02	61.81	37.04	52.31	58.68	**88.89**
Sub12	57.41	51.85	59.03	55.79	42.59	51.39	64.81	**66.9**
Sub13	53.7	53.7	58.33	54.86	40.74	54.63	68.29	**72.22**
Sub14	64.81	70.37	51.39	63.19	40.74	58.8	76.39	**76.85**
Sub15	**62.5**	45.83	54.17	**62.5**	45.83	59.38	48.44	**62.5**
Sub16	50	50	51.16	50.69	38.89	46.3	43.87	**59.26**
Sub17	66.67	66.67	62.73	62.96	37.04	62.96	74.07	**81.71**
Sub18	57.41	42.59	63.43	62.73	42.59	54.63	52.31	**78.7**
Sub19	46.3	50	53.94	57.18	46.3	**62.96**	45.95	**62.96**
Sub20	46.3	50	**60.42**	57.64	42.59	50.93	49.07	**60.42**
AVG	51.46	52.11	54.95	56.76*	41.83**	54.13	54.85	**69.20****

TABLE 5. The t-test results between the proposed method and each Other method.

	CSP	FBCSP	MCNN	MSCNN	TD-Atten	MTFB-CNN	EEGNet
Proposed	6.68e-08**	1.72e-07**	5.14e-07**	7.53e-06**	3.63e-15**	1.673e-07**	3.113e-05**

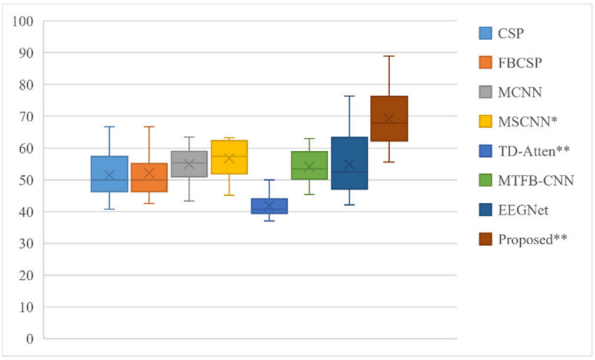


FIGURE 5. Comparison of box plots of the correctness of each method. The 'x' mark indicates the average accuracy of the method, while significance differences are denoted in the legend (represented as $*$ for $p < 0.05$ and $**$ for $p < 0.01$). The classification accuracy rate of deep learning is generally higher than traditional machine learning methods. Compared with other selected methods, the method proposed in this paper significantly improves classification accuracy.

as shown in Fig. 6. Participants 6, 7, 11, and 14 exhibited relatively uniform classification results across various motor imagery tasks, showing no apparent bias. In contrast, participants 3 and 8 demonstrated distinctive differences in their performance across different motor imagery tasks. Subject 3 achieved better results in the imagination of fist clenching and shoulder rotation, but it became indistinguishable in elbow flexion, while subject 8 was indistinguishable in shoulder rotation. Fig. 7 depicts the confusion matrix of the classification results for all participants. From the perspective of all participant data, the classifier proposed in this study does not exhibit significant bias. It only slightly underperforms compared to the other two categories in the task of elbow flexion classification.

D. DATA AUGMENTATION RESULTS

To investigate the impact of data augmentation on the experimental results, comparative experiments were conducted, and the outcomes are presented in Table 7. In this section of the experiments, a 4 times data augmentation was employed, equivalent to down sampling each data segment from the

TABLE 6. Comparison of parameters and results of each method.

Method	Year	Accuracy (%)	Parameters Counts	kappa	f1
MCNN	2019	54.95	34,496,137	0.30	0.55
MSCNN	2022	56.76	423,544,675	0.33	0.57
TD-Atten	2022	41.83	66,547	0.01	0.42
MTFB-CNN	2023	54.13	334,692	0.27	0.54
Proposed		69.20	38,543	0.51	0.67

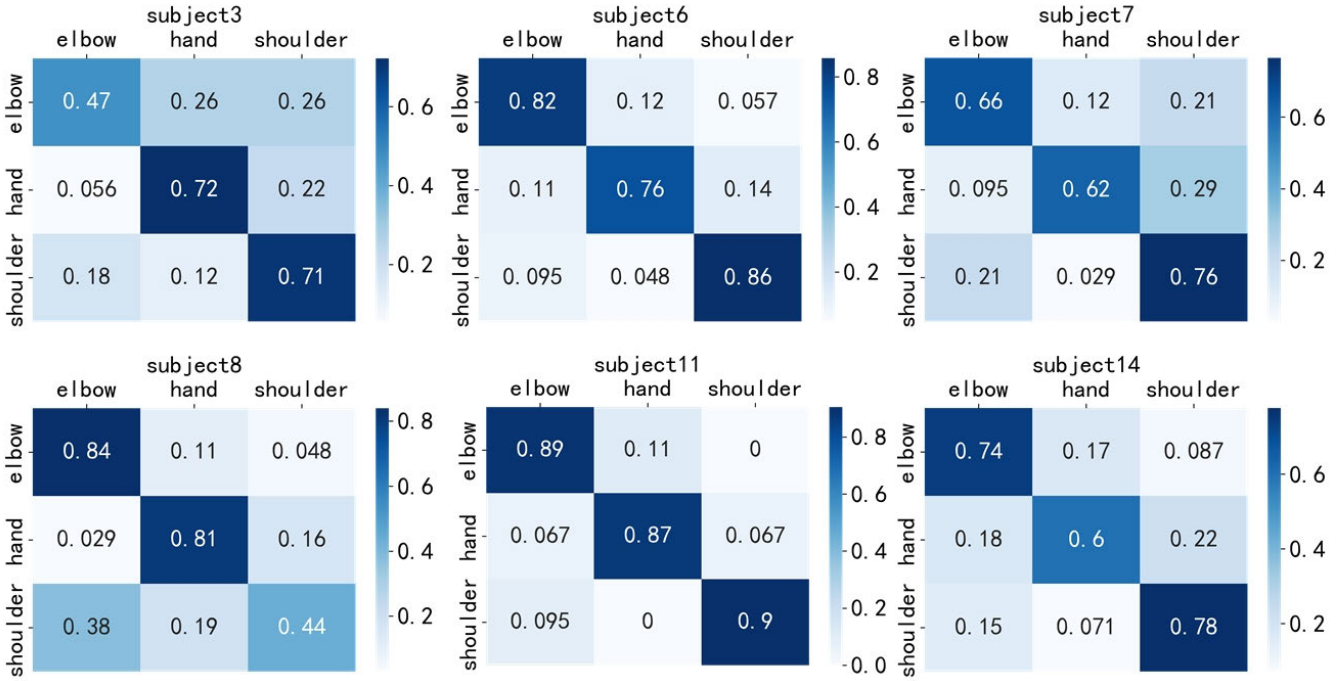


FIGURE 6. Confusion matrix for partial subject classification. Subjects 6, 7, 11, and 14 had relatively average classification rates on each action. However, subject 3 became indistinguishable in the item of elbow flexion. Subject 8 became indistinguishable in the item of shoulder rotation.

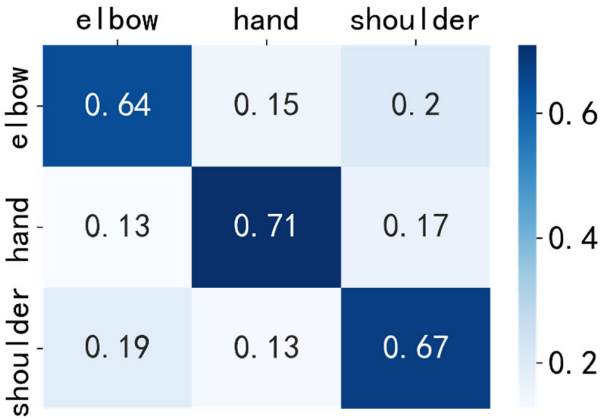


FIGURE 7. An overall confusion matrix with data from all of the subjects.

original 1000Hz sampling rate to 250Hz, resulting in a four-fold increase in the overall dataset size. As indicated in Table 7, after data augmentation, there was a noticeable improvement in the accuracy of most participants' data, reflecting the impact of data augmentation.

E. ABLATION EXPERIMENTS

In order to study the influence of each module in the deep learning model on the final classification result, an ablation experiment was designed. The model proposed in this paper is based on the multi-band EEG signal and the attention mechanism between frequency bands. Therefore, the ablation experiment mainly focuses on the multi-band EEG signal and the attention mechanism module. The first experiment represents the most primitive form of the model, which is based on a single frequency band without incorporating attention mechanisms. In this experimental process, we opted for the most prevalent MI frequency band (8∼30Hz) [14]. Other preprocessing operations remained consistent with those previously mentioned. The second experiment introduces multi-frequency EEG signals but does not involve any attention mechanisms. The frequency bands of the data used in this experiment match the frequency bands selected for the final model (i.e., 1-4Hz, 4-8Hz, 8-13Hz, 13-31Hz, 31-40Hz). The third experiment encompasses both multi-frequency signals and inter-frequency attention mechanisms.

TABLE 7. Comparison of classification results between original data and augmented data (in percentage).

	original data	augmented data		original data	augmented data
Sub1	46.3	**55.56**	Sub11	77.61	**88.89**
Sub2	**63.24**	56.94	Sub12	63.89	**66.9**
Sub3	60.19	**62.96**	Sub13	**73.13**	72.22
Sub4	68.52	**74.54**	Sub14	59.26	**76.85**
Sub5	52.94	**62.04**	Sub15	50	**62.5**
Sub6	64.18	**80.79**	Sub16	48.53	**59.26**
Sub7	73.53	**74.07**	Sub17	70.15	**81.71**
Sub8	**71.3**	71.06	Sub18	73.15	**78.7**
Sub9	64.71	**66.9**	Sub19	**65.67**	62.96
Sub10	55.22	**68.75**	Sub20	60.19	**60.42**

TABLE 8. Results of ablation experiments (denoted as $*$ for $p < 0.05$ and $**$ for $p < 0.01$).

Multi-Freq	Atten	Accuracy (%)	kappa	f1
-	-	54.85	0.32	0.55
✓	-	65.15**	0.47	0.65
✓	✓	69.20**	0.51	0.67

In the three stages of the ablation experiment, the parameters and methods used in model training remained consistent, and the results are shown in Table 8. When utilizing the model based on a single frequency band, an average accuracy of 54.85% was achieved, which was lower than the results of the other two experiments. Subsequently, when transitioning to the model with multiple frequency bands, the accuracy increased by 10.3%. This demonstrates that the incorporation of multi-frequency signals significantly enhances the discriminability of fine motor imagery EEG signals. Furthermore, when the inter-frequency attention mechanism was integrated on top of the multi-frequency signals, the accuracy saw an additional improvement of 4.05%.

Utilizing the network as a baseline classification method without multiple frequency bands and attention mechanisms, a t-test was conducted. As indicated in Table 8, there exists a significant difference in accuracy when comparing the models with multiple frequency bands and attention mechanisms to the baseline method ($p < 0.01$). Therefore, we can conclude that the attention modules designed for different frequency bands are effective in assisting the network to better learn the features of FMI-EEG signals, thereby improving classification accuracy.

F. ERS/ERD ANALYSIS

Previous research has demonstrated that the C3 electrode exhibits the most pronounced Event-Related Desynchronization (ERD) response in comparison to other electrodes [33]. Therefore, we selected a participant, utilized the C3 electrode, and performed multi-segment averaging of EEG data. We considered the data before time point zero (i.e., data from -1 to 0 seconds) as baseline data. Subsequently, we plotted comparative graphs illustrating the energy variations of different FMI actions across various frequency bands [34], as depicted in Fig. 8. It is observable from the figure that a distinct ERD phenomenon is evident at 0.5 seconds following the appearance of the FMI stimulus, as highlighted within the green box. Additionally, the ERD phenomenon for the fist-clench FMI action by the subject (represented by the red curve) is notably weaker when compared to the other two FMI actions. Based on Fig. 8 and the work of other researchers [35], [36], the 8∼13Hz frequency band, exhibiting the most pronounced ERD phenomenon, was selected. The average energy distribution changes of this participant were mapped onto a topographical representation, as shown in Fig. 9. In the three motor imagery tasks of elbow flexion, fist clenching, and shoulder rotation, the contralateral motor area was significantly activated, and the ipsilateral motor area was also activated. However, it was not as apparent as the contralateral side.

After averaging the MI data of all participants, comparative graphs illustrating the energy variations of different FMI actions across various frequency bands were created, using data before 0 seconds as the baseline, as shown in Fig. 10. A noticeable ERD phenomenon appeared at 0.5 seconds after the onset of FMI (highlighted in green boxes in the figure). In the beta frequency band, there was a slight rebound following a substantial ERD. Therefore, we segmented the beta band to investigate this occurrence. It was observed that after the signal rebound, there were one or more minor subsequent declines (highlighted in purple boxes in the figure), and higher-frequency signals exhibited more distinct patterns. Similarly, based on Fig. 10 and the work of other researchers [35], [36], the 8∼13Hz frequency band, showing the most prominent ERD phenomenon, was selected to generate average power change distributions for all participant data, as depicted in Fig. 11. All three types of FMI signals demonstrated significant activation in the contralateral motor area, notably stronger than the ipsilateral side.

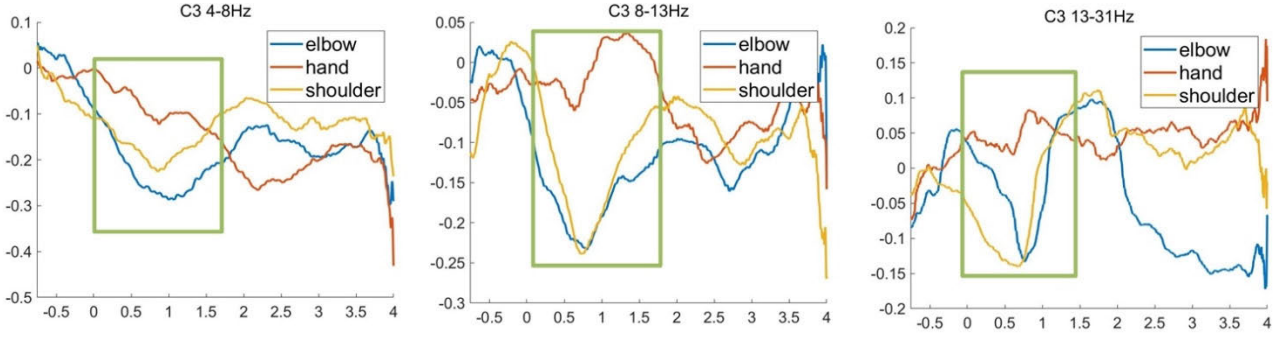


FIGURE 8. C3 electrode, comparison chart of energy changes of different actions in different frequency bands for a selected participant. The vertical axis represents the change in power relative to the baseline (%) and the horizontal axis depicts the time (s) for motor imagery. The green box indicates the occurrence of ERD phenomenon.

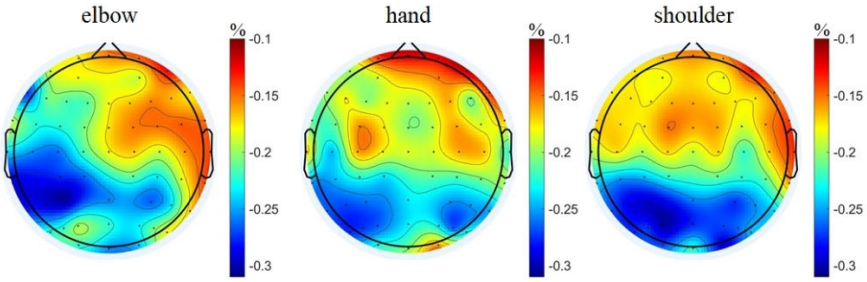


FIGURE 9. Topographical map of energy changes distribution for different motor imagination movements for a selected participant. The activation of the motor area on the contralateral side of the three MIs is relatively apparent, with slight activation also observed in the ipsilateral motor area, albeit less distinct compared to the contralateral side.

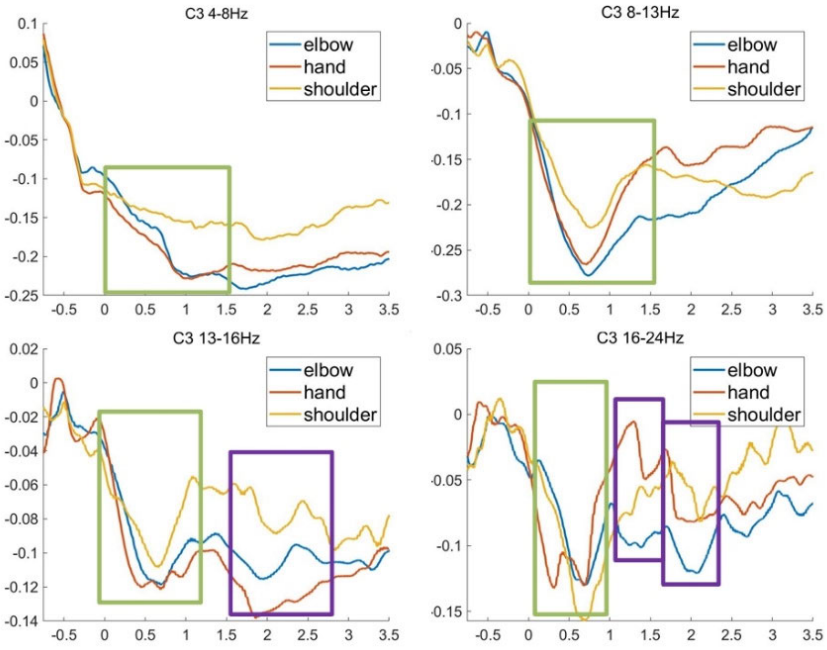


FIGURE 10. C3 electrode, comparison chart of energy changes of different actions in different frequency bands for all participants. The vertical axis represents the change in power relative to the baseline (%), while the horizontal axis indicates the time (s) during motor imagery. The green box indicates the occurrence of ERD phenomenon.

For the ERD phenomena across different frequency bands and categories, pairwise t-tests were conducted and the results are presented in Table 9. The findings indicate significant differences ($p < 0.05$) in ERD phenomena between different categories within distinct frequency bands. This suggests that, overall, the EEG signals related to these three types of motor

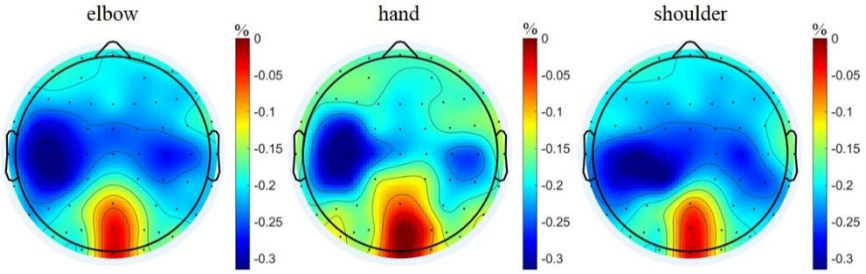


FIGURE 11. Topographical distribution of average band power changes for all participants.

TABLE 9. The t-test results for ERD phenomena between different categories.

category	*p* value			
	4-8 Hz	8-13 Hz	13-16 Hz	16-24 Hz
elbow VS hand	3.53838E-09	2.84785E-30	1.0889E-34	2.06612E-30
elbow VS shoulder	1.36328E-57	2.9033E-169	8.2856E-22	1.88535E-20
hand VS shoulder	6.4502E-225	1.07448E-66	7.92368E-77	3.76274E-39

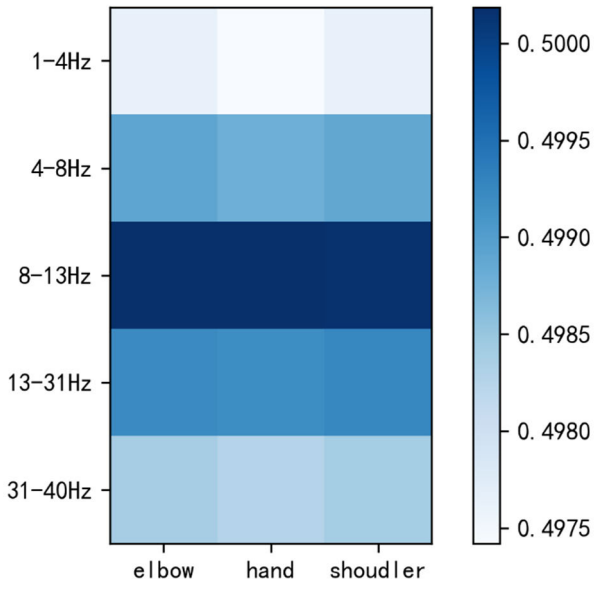


FIGURE 12. Visualization of attention network output weights for different frequency bands and categories of EEG signals.

imagery are distinguishable, and there are features recognizable by the classifier across various frequency bands.

G. VISUALIZATION OF FREQUENCY BAND WEIGHTS

Different frequency bands play distinct roles in the classification of motor imagery. To investigate the impact of different frequency bands on classifier performance, the weights of the frequency band attention network were exported and visualized, as depicted in Fig. 12. The horizontal axis represents different categories of motor imagery, while the vertical axis represents different frequency bands, and the color intensity indicates the weight magnitude. From Fig. 12, it is evident that the 8-13Hz frequency band of EEG signals exhibits the highest contribution, followed by the 4-8Hz and 13-31Hz frequency bands. This observation aligns with the prominent ERD phenomena in several frequency bands as illustrated in Fig. 10. Consequently, the attention network assigns different weights to EEG signals from different frequency bands and categories, aiding the classifier in extracting features from distinct frequency ranges. This reaffirms that different frequency bands make varied contributions to motor imagery classification.

V. DISCUSSION

In this study, different movements of the same upper limb were used as an MI paradigm called the FMI paradigm. A multi-band attention mechanism and feature extraction neural network decoded and classified FMI-EEG data. It is shown that this network outperforms other comparative methods regarding classification results and the number of network parameters. In the triple classification task, the best classification result of this method is 88.89%, and the average of all subjects is 69.2%. Compared with the classic machine learning method, for the CSP+SVM method, the classification accuracy rate has increased by 17.74%, and for the FBCSP+SVM method, the accuracy rate has increased by 17.09%. As can be seen from Table 6, compared with the four selected deep learning methods [11], [16], [17], [32], the accuracy of the proposed method has been significantly improved, with the highest improvement of 23.37%. At the same time, this model has significantly fewer trainable parameters than other models, which can make the model converge faster during training, consume less computing time when performing classification tasks, and save more GPU resources.

Regarding the design of the paradigm flow, prior to the formal data collection experiment, a small group of individuals was recruited for an investigation into the paradigm

flow and timing. The majority indicated that the time of particularly focused attention during the three categories of motor imagery experiments is approximately 20 minutes, with the need for intermittent rest periods. Consequently, this study sets the duration of one session at around 20 minutes, segmented into several blocks with designated rest intervals. In order to obtain an adequate amount of data for subsequent analysis, and with the participants' consent, the experiment will consist of a total of 3 sessions. Each session will include sufficient rest time to allow participants to recuperate, ensuring the quality of the collected data.

In this study, the utilization of down-sampling as a data augmentation technique is aimed at expanding the limited dataset without sacrificing any data. Additionally, due to hardware constraints, it was not feasible to load all data into memory and GPU memory simultaneously for data processing and model training. Consequently, frequent disk I/O operations during training consumed significant time. Down sampling the data led to a reduction in the model's parameter count, allowing us to read more data from the disk in each operation, reducing I/O operations and expediting training time. Furthermore, a smaller model size implies lower hardware requirements, facilitating easier development and deployment of future BCI systems.

The attention mechanism between frequency bands can help the network to learn independently, assign weight to data in different frequency bands, and improve classification accuracy. The feature extraction module can perform further feature extraction for the time, space, and frequency domains [29]. Therefore, prior to feature extraction, we manually divided the data into five frequency bands, hoping that the attention mechanism could globally weigh these different frequency bands. After weighting, the feature extraction module targeted these frequency bands for specific feature extraction, thereby extracting more detailed features. The ablation experiment results in Table 8 show the effectiveness of each module.

In both Fig. 8 and Fig. 10, the signal rebound phenomenon around 1 second is possibly associated with our experimental protocol for the participants. Similar occurrences of this phenomenon have also appeared in the works of other relevant researchers [35], [36]. We instructed each participant to repeatedly imagine the action during every single FMI process. However, the execution frequency varied among participants, leading to a superposition of misaligned ERD/ERS signals across participants, resulting in this observed phenomenon. In subsequent experiments, we will instruct participants to perform repeated motor imagery at a fixed frequency (e.g., 2Hz) to avoid the occurrence of such situations. From Fig. 8, it's evident that the fist-clenching ERD phenomenon for this participant is notably weaker than the other two actions. However, Fig. 10 doesn't indicate a specific FMI action with significantly weaker ERD compared to the others. Additionally, we observed in Fig. 9 that the ERD occurrence area for this participant is positioned further back than the typically recognized motor area. However, in the brain average power distribution of all participants shown in Fig. 11, this observation did not manifest, in line with the expected outcomes [11], [35]. Hence, we believe these occurrences are due to individual differences, which are subdued in the average of all participant data.

Brain-computer interface relies on the function of the normal brain to a certain extent. Patients with cortical injuries (such as ALS and stroke) may affect the use of BCI systems, and the ERD pattern of paralyzed patients during motor imagination is similar to that of non-disabled people [37]. At the same time, some studies have shown that ERD in the motor rhythm of brain electrical signals plays an important role in the motor recovery of stroke patients [38], and stroke patients can reach their maximum exercise level through training [39]. Therefore, in the future, this FMI decoding model can be applied to patients to help them restore upper limb motor functions.

In the context of using the unimanual motor imagery paradigm, research has indicated that patients with gradually increasing ERD in the lesioned hemisphere, as compared to the healthy hemisphere, exhibit greater improvements in motor function. Furthermore, there exists a stronger correlation between the improvement in motor function and the ERD activity in the lesioned hemisphere [38]. Consequently, it is advisable to prioritize the training of motor imagery capabilities in the lesioned hemisphere, in order to facilitate maximal recovery of motor function in patients.

Currently, there are still some limitations in the research on the decoding of imagined fine motor movements of the unilateral upper limb. Due to various factors such as psychological or physiological, there is substantial variability in the electroencephalographic signals of each individual [40], and individual differences become even more pronounced in task-related states [41]. Consequently, it is possible that some participants may exhibit poorer performance on specific classification tasks. Although some methods can compensate for this inter-subject variation to some extent, it still needs a more robust understanding of the changes in the distribution of EEG signal characteristics across subjects and sessions [42]. This study primarily focused on the investigation of decoding models tailored for individual participants, and the robustness needs improvement.

In the selection of experimental data for this study, 20 healthy individuals were utilized as subjects, and no FMI-EEG data from stroke patients were collected. In subsequent research, it is imperative to collect EEG data from actual stroke patients to validate the genuine applicability of the system. It is important to note that stroke patients differ from healthy individuals. Most patients struggle with prolonged attention span and must consider potential safety hazards when exposed to external stimuli. Therefore, when collecting FMI-EEG data from stroke patients, certain modifications to the paradigm workflow are necessary. Additionally, constant attention to the patient's state during the experimental process is crucial to prevent unexpected occurrences.

Moreover, concerning model training, considering the utilization of transfer learning to enhance the robustness of the network model is contemplated. This approach aims to diminish the training time for new participants while mitigating the impact of poor performance on specific classification tasks arising from individual differences among participants. Furthermore, the utilization of transfer learning is aimed at enabling the classification model to leverage the knowledge acquired from data obtained from healthy individuals to address the FMI-EEG classification challenges in real patients. In the aspect of data acquisition, this study opted for EEG as the primary source of raw signals. Subsequent research endeavors will contemplate incorporating additional signals from kinematics or neuroimaging techniques to supplement the analysis.

VI. CONCLUSION

Classic motor imagery, which makes distinctions between different limbs, can no longer meet hemiplegic patients' upper limb rehabilitation needs. Therefore, this study designed three paradigms of FMI movement (elbow, shoulder, and hand) for unilateral upper limbs. Moreover, 20 healthy subjects were recruited and the Neuracle's wireless amplifier was used to collect EEG signals and to study the EEG decoding of FMI.

The attention mechanism for the frequency band and the deep learning neural network were used for feature extraction, and the performance in the triple classification task reached an average of 69.2%. The ablation experiments demonstrate the necessity of each module. These results demonstrate the feasibility of the multi-band attention mechanism for FMI-EEG data decoding and the potential of this FMI paradigm for exoskeleton control, providing a new training method for upper limb rehabilitation.

DECLARATION OF INTEREST

None

REFERENCES

[1] N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "EEG-based brain-computer interfaces using motor-imagery: Techniques and challenges," *Sensors*, vol. 19, no. 6, p. 1423, Mar. 2019.

[2] Y. Hou, T. Chen, X. Lun, and F. Wang, "A novel method for classification of multi-class motor imagery tasks based on feature fusion," *Neurosci. Res.*, vol. 176, pp. 40–48, Mar. 2022.

[3] N. Marotta, A. Ammendolia, C. Marinaro, A. Demeco, L. Moggio, and C. Costantino, "International classification of functioning, disability and health (ICF) and correlation between disability and finance assets in chronic stroke patients," *Acta Bio Medica, Atenei Parmensis*, vol. 91, no. 3, Jan. 2020, Art. no. e2020064.

[4] M. A. Khan, R. Das, H. K. Iversen, and S. Puthusserypady, "Review on motor imagery based BCI systems for upper limb post-stroke neurorehabilitation: From designing to application," *Comput. Biol. Med.*, vol. 123, Aug. 2020, Art. no. 103843.

[5] T. C. Machado, A. A. Carregosa, M. S. Santos, N. M. D. S. Ribeiro, and A. Melo, "Efficacy of motor imagery additional to motor-based therapy in the recovery of motor function of the upper limb in post-stroke individuals: A systematic review," *Topics Stroke Rehabil.*, vol. 26, no. 7, pp. 548–553, Oct. 2019.

[6] E. M. D. Santos, C. A. Fernandes, and G. Castellano, "Performance of stroke patients using a brain-computer interface during motor imagery: A systematic review," *Res. Biomed. Eng.*, vol. 39, no. 2, pp. 451–465, May 2023.

[7] B. E. Olivas-Padilla and M. I. Chacon-Murguia, "Classification of multiple motor imagery using deep convolutional neural networks and spatial filters," *Appl. Soft Comput.*, vol. 75, pp. 461–472, Feb. 2019.

[8] Y. Miao, J. Jin, I. Daly, C. Zuo, X. Wang, A. Cichocki, and T.-P. Jung, "Learning common time-frequency-spatial patterns for motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 699–707, 2021.

[9] P. Ofner, A. Schwarz, J. Pereira, and G. R. Müller-Putz, "Upper limb movements can be decoded from the time-domain of low-frequency EEG," *PLoS ONE*, vol. 12, no. 8, Aug. 2017, Art. no. e0182578.

[10] A. Suwannarat, S. Pan-Ngum, and P. Israsena, "Comparison of EEG measurement of upper limb movement in motor imagery training system," *Biomed. Eng. OnLine*, vol. 17, no. 1, p. 103, Dec. 2018.

[11] X. Ma, S. Qiu, and H. He, "Time-distributed attention network for EEG-based motor imagery decoding from the same limb," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 496–508, 2022.

[12] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000.

[13] A. Palumbo, V. Gramigna, B. Calabrese, and N. Ielpo, "Motor-imagery EEG-based BCIs in wheelchair movement and control: A systematic literature review," *Sensors*, vol. 21, no. 18, p. 6285, Sep. 2021.

[14] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Deep learning for motor imagery EEG-based classification: A review," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102172.

[15] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals," *Comput. Biol. Med.*, vol. 143, Apr. 2022, Art. no. 105288.

[16] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. S. Hossain, "Deep learning for EEG motor imagery classification based on multi-layer CNNs feature fusion," *Future Gener. Comput. Syst.*, vol. 101, pp. 542–554, Dec. 2019.

[17] A. M. Roy, "An efficient multi-scale CNN model with intrinsic feature integration for motor imagery EEG subject classification in brain-machine interfaces," *Biomed. Signal Process. Control*, vol. 74, Apr. 2022, Art. no. 103496.

[18] Y. Zhang, S. Qiu, and H. He, "Multimodal motor imagery decoding method based on temporal spatial feature alignment and fusion," *J. Neural Eng.*, vol. 20, no. 2, Apr. 2023, Art. no. 026009.

[19] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 4–14, Jan. 2016.

[20] B. J. Edelman, J. Meng, D. Suma, C. Zurn, E. Nagarajan, B. S. Baxter, C. C. Cline, and B. He, "Noninvasive neuroimaging enhances continuous neural tracking for robotic device control," *Sci. Robot.*, vol. 4, no. 31, Jun. 2019, Art. no. eaaw6844.

[21] L. Yang, T. Shi, J. Lv, Y. Liu, Y. Dai, and L. Zou, "A multi-feature fusion decoding study for unilateral upper-limb fine motor imagery," *Math. Biosci. Eng.*, vol. 20, no. 2, pp. 2482–2500, 2022.

[22] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-based EEG classification in motor imagery tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, Nov. 2018.

[23] J. Chen, W. Yi, D. Wang, J. Du, L. Fu, and T. Li, "FB-CGANet: Filter bank channel group attention network for multi-class motor imagery classification," *J. Neural Eng.*, vol. 19, no. 1, Jan. 2022, Art. no. 016011.

[24] D. W. Prabowo, H. A. Nugroho, N. A. Setiawan, and J. Debayle, "A systematic literature review of emotion recognition using EEG signals," *Cognit. Syst. Res.*, vol. 82, Dec. 2023, Art. no. 101152.

[25] K. M. Hossain, M. A. Islam, S. Hossain, A. Nijholt, and M. A. R. Ahad, "Status of deep learning for EEG-based brain-computer interface applications," *Frontiers Comput. Neurosci.*, vol. 16, Jan. 2023, Art. no. 1006763.

[26] B. Abibullaev, A. Keutayeva, and A. Zollanvari, "Deep learning in EEG-based BCIs: A comprehensive review of transformer models, advantages, challenges, and applications," *IEEE Access*, vol. 11, pp. 127271–127301, 2023.

[27] A. Vaswani, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–7.
[28] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
[29] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.
[30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[31] K. K. Ang, Z. Yang Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 2390–2397.
[32] H. Li, H. Chen, Z. Jia, R. Zhang, and F. Yin, "A parallel multi-scale time-frequency block convolutional neural network based on channel attention module for motor imagery classification," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104066.
[33] L. Yakovlev, N. Syrov, A. Miroshnikov, M. Lebedev, and A. Kaplan, "Event-related desynchronization induced by tactile imagery: An EEG study," *Eneuro*, vol. 10, no. 6, p. 455, Jun. 2023.
[34] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov. 1999.
[35] Y. Chu, X. Zhao, Y. Zou, W. Xu, G. Song, J. Han, and Y. Zhao, "Decoding multiclass motor imagery EEG from the same upper limb by combining Riemannian geometry features and partial least squares regression," *J. Neural Eng.*, vol. 17, no. 4, Aug. 2020, Art. no. 046029.
[36] B. Xu, Z. Wei, A. Song, C. Wu, D. Zhang, W. Li, G. Xu, H. Li, and H. Zeng, "Phase synchronization information for classifying motor imagery EEG from the same limb," *IEEE Access*, vol. 7, pp. 153842–153852, 2019.
[37] S. Machado, L. F. Almada, and R. N. Annavarapu, "Progress and prospects in EEG-based brain-computer interface: Clinical applications in neurorehabilitation," *J. Rehabil.*, vol. 1, no. 1, p. 29, 2013.
[38] A. M. Ray, T. D. C. Figueiredo, E. López-Larraz, N. Birbaumer, and A. Ramos-Murguialday, "Brain oscillatory activity as a biomarker of motor recovery in chronic stroke," *Human Brain Mapping*, vol. 41, no. 5, pp. 1296–1308, Apr. 2020.
[39] C. B. Tabernig, L. C. Carrere, C. A. Lopez, and C. Ballario, "EEG event-related desynchronization of patients with stroke during motor imagery of hand movement," *J. Phys., Conf. Ser.*, vol. 705, Apr. 2016, Art. no. 012059.
[40] K. Zhang, G. Xu, L. Chen, P. Tian, C. Han, S. Zhang, and N. Duan, "Instance transfer subject-dependent strategy for motor imagery signal classification using deep convolutional neural networks," *Comput. Math. Methods Med.*, vol. 2020, pp. 1–10, Aug. 2020.
[41] E. Gibson, N. J. Lobaugh, S. Joordens, and A. R. McIntosh, "EEG variability: Task-driven or subject-driven signal of interest?" *NeuroImage*, vol. 252, May 2022, Art. no. 119034.
[42] G. Huang, Z. Zhao, S. Zhang, Z. Hu, J. Fan, M. Fu, J. Chen, Y. Xiao, J. Wang, and G. Dan, "Discrepancy between inter- and intra-subject variability in EEG-based motor imagery brain-computer interface: Evidence from multiple perspectives," *Frontiers Neurosci.*, vol. 17, Feb. 2023, Art. no. 1122661.

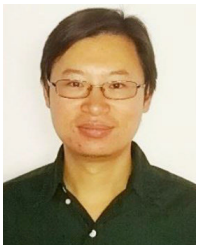
TIANYU SHI is currently pursuing the M.S. degree in electronic science and technology with Changzhou University, Changzhou, China. His current research interest includes brain–computer interface.

XIANG GU is currently pursuing the master's degree with the School of Computer Science and Artificial Intelligence, Changzhou University, China. His current research interest includes bioelectrical signal processing.

HUI BI received the Ph.D. degree in computer science and technology from Southeast University, in 2018. She is currently a Lecturer with the School of Aliyun Big Data, Changzhou University. Her current research interests include algorithm analysis and design, signal, and image processing.

JIDONG LV received the Ph.D. degree from Jiangsu University, Zhenjiang, China, in 2012. He is currently with Changzhou University. His current research interests include bioelectrical signal acquisition and processing.

YAN LIU was born in 1984. She received the Ph.D. degree from Harbin Institute of Technology, Harbin, China, in 2013. Her current research interests include EEG/MEG/MRI processing and epilepsy seizure prediction and foci localization.



YAKANG DAI was born in 1982. He received the Ph.D. degree from Chinese Academy of Sciences, Beijing, China, in 2009. His current research interests include medical image analysis and BCI.

LING ZOU is currently a Professor with the School of Microelectronics and Control Engineering, Changzhou University, China. Her research interests include neural engineering, brain–computer interface, and biomedical signal processing.