Decoding Multi-Class EEG Signals of Hand Movement Using Multivariate Empirical Mode Decomposition and Convolutional Neural Network

Yi Tao, Weiwei Xu, Guangming Wang, Ziwen Yuan, Maode Wang, Michael Houston, Yingchun Zhang, *Senior Member, IEEE*, Badong Chen, Xiangguo Yan, and Gang Wang, *Member, IEEE*

***Abstract*— Brain-computer interface (BCI) is a technology that connects the human brain and external devices. Many studies have shown the possibility of using it to restore motor control in stroke patients. One specific challenge of such BCI is that the classification accuracy is not high enough for multi-class movements. In this study, by using Multivariate Empirical Mode Decomposition (MEMD) and Convolutional Neural Network (CNN), a novel algorithm (MECN) was proposed to decode EEG signals for four kinds of hand movements. Firstly, the MEMD was used to decompose the movement-related electroencephalogram (EEG) signals to obtain the multivariate intrinsic empirical functions (MIMFs). Then, the optimal MIMFs fusion was performed based on sequential forward selection algorithm. Finally, the selected MIMFs were input to the CNN model for discriminating four kinds of hand movements. The average classification accuracy of thirteen subjects over the six-fold cross-validation reached 81.14% for 2s-data before the movement onset and 81.08% for 2s-data after the movement onset. The MECN method achieved statistically significant improvement on the state-of-the-art methods. The results showed that the algorithm proposed in this study can effectively decode four kinds of hand movements based on EEG signals.**

***Index Terms*— Electroencephalogram, hand movement, brain–computer interface, multivariate empirical mode decomposition, convolutional neural network.**

Manuscript received 16 March 2022; revised 3 July 2022 and 15 September 2022; accepted 19 September 2022. Date of publication 22 September 2022; date of current version 30 September 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 32071372, Grant 62271385, and Grant 61471291; in part by the Science and Technology Project of Sichuan under Grant 2022YFS0030; and in part by the Natural Science Basic Research Program of Shaanxi through the Program under Grant 2020JM-037. *(Corresponding author: Gang Wang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of Xi'an Jiaotong University Health Science Center under Application No. 2020-620.

Yi Tao, Weiwei Xu, Guangming Wang, Xiangguo Yan, and Gang Wang are with the Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049, China, also with the National Engineering Research Center for Healthcare Devices, Guangzhou 510500, China, and also with the Key Laboratory of Neuro-Informatics and Rehabilitation Engineering of Ministry of Civil Affairs, Xi'an 710049, China (e-mail: ggwang@xjtu.edu.cn).

Ziwen Yuan is with the Department of Rehabilitation, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, China, and also with the Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049, China.

Maode Wang is with the Department of Neurosurgery, First Affiliated Hospital, Xi'an Jiaotong University, Xi'an 710061, China.

Michael Houston and Yingchun Zhang are with the Department of Biomedical Engineering, University of Houston, Houston, TX 77204 USA.

Badong Chen is with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China.

Digital Object Identifier 10.1109/TNSRE.2022.3208710

I. Introduction

BRAIN computer interface (BCI) is a technology that establishes a communication system between the human brain and external devices [1]. It has been applied in many fields, such as stroke rehabilitation [2], [3], prosthetic control [4], quadcopter control [5], speech synthesis [6], and emotion recognition [7]. The techniques aiming to reconstruct hand motor function have been extensively studied [8], [9], [10] and are expected to enable rehabilitation training for stroke patients. It's proved that the task variability can improve the performance in the retention session of learned motor skills and increase the generalization of learning to new skills [11]. Hence, decoding fine hand movement intention is of great value for rehabilitation training of hand function.

As is well-known, the recognition performance of hand movements using surface electromyographic (sEMG) signals or angle sensor data can achieve an accuracy of closed to 100% in some previous studies [12], [13], [14], [15], [16], [17]. However, for some severe stroke patients, they may not have sufficiently high levels of muscle activity and are hard to achieve normal hand movements of daily living. Then, the sEMG signals or angle sensor data cannot provide the adequate information to decode different hand movements. Hence, the EEG signals were used to classify different hand movements instead of sEMG signals and angle sensor data.

Many EEG-based classification systems work by decoding motor imagery (MI) signals from different parts of human body, such as right/left hand, feet, and tongue [18], [19], [20], regardless of the actual output command [21], because these tasks activate various areas of the cerebral cortex which are distant enough to be well distinguished [22], [23], [24]. Nevertheless, these tasks are inconsistent with actual instructions in most cases, and sustained motor imagery is not natural neither comfortable for the user. By contrast, the motor execution (ME) tasks are consistent with the rehabilitation device feedback, and the BCI can better restore motor function of patients by inducing activity-dependent brain plasticity [25]. What's more, comparing with MI, the ME produces high movement-related brain activity to accurately classify different hand movements. References [26], [27] Therefore, the EEG of ME tasks were used to decode the movement intentions in this study.

The typical application scenarios of decoding multi-class EEG signals of hand movements is the rehabilitation training of hand function. However, since the stroke patients have insufficient motor capability to perform normal hand movements in daily life, they are unable to successfully achieve the ME experiment paradigm. On the other hand, because the motor cortex of stroke patients has suffered from injuries, the strength of EEG signals related to hand movements becomes weak and the EEG pattern of brain activities among different hand movements is also difficult to be discriminated. Thus, the neurologically healthy subjects with ME tasks were used to verify the feasibility of decoding multi-class movements from same hand. Besides, the EEG classification algorithm based on healthy subjects with ME tasks would be applied to the hand movement decoding of stroke patients by transfer learning in the future. In order to accomplish the objective of multi-class movement pattern recognition of same hand with stroke patients, many attempts have been firstly made to focus on the EEG classification of healthy subjects [28], [29], [30], [31]. Consequently, the neurologically healthy subjects were tested for our proposed method in this study.

Most studies have attempted to extract different movement-related features, such as kinds of temporal and spectral features [32], [33], [34], [35], [36], [37], [38], phase-lock-based features [39], time-frequency (TF) map features of each source signal in the motor cortex [40]. Some other studies proposed to use different neural network structures for feature extraction and classifying, such as compact convolutional neural network named EEGNet [41], deep ConvNets [42]. Previous studies have achieved good results in distinguishing 2-class hand movement of one hand, with a classification accuracy of around 80%. However, the accuracy of multi-class hand movements of one hand was between 50% and 70%, which was not high enough to satisfy the application of rehabilitation training at present [43]. Therefore, in this study, a novel algorithm (MECN) using Multivariate Empirical Mode Decomposition (MEMD) and Convolutional Neural Network (CNN) was proposed to classify four kinds of hand movement. The algorithm was performed on the recorded signals before and after movement onset to decode the EEG signals. The remainder of this paper is organized as follows: Section II introduces the experiment and the proposed algorithm in detail. Section III presents the results of the proposed algorithm on the collected data. Section IV describes the discussions. Section V concludes this paper.

II. Methods

A. Experimental Protocol and Data Acquisition

Sixteen subjects (1 female and 15 males: 23.1 $\pm$ 2.6 years old) participated in this experiment. They are neurologically healthy and right-handed. Written informed consent was obtained from all subjects before the experiments and the study was approved the 26 March 2020 by the Institutional Review Board at Xi'an Jiaotong University, China, ref. 2020-620. Due to the poor quality of EEG or EMG, the data of three subjects were excluded from further analysis (1 female and 2 males).

EEG experiments were carried out in a quiet room. The subjects were seated in a comfortable chair with their right arm resting on the table in front of them and they were approximately 0.5 m distance from the screen. The experimental session of each movement task consisted of 3 blocks, as shown in Fig. 1. At the beginning of each block, a five-second window on the computer screen with words *Start* indicated that this block was about to begin, and subjects were reminded to prepare for the coming trials. One trial included a four-second cue part with words *Ready for Tip Pinch* (or *Multiple Tip Pinch*, *Hand Close*, *Hand Open*), a four-second execution part instructed subjects to execute the movement displayed on the screen and a six-second rest part in which subjects were allowed to relax, blink, and swallow. Each subject performed 60 trials in a session, with a 5-min rest between each block of 20 trials.

EEG signals were recorded from 30 scalp electrodes and EMG signals were recorded simultaneously from 4 Ag/AgCl electrodes (Neuroscan Systems, Compumedics, Charlotte, NC, USA). The EEG electrodes were placed at FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T7, C3, CZ, C4, T8, TP7, CP3, CPZ, CP4, TP8, P7, P3, PZ, P4, P8, O1, OZ, O2 locations according to international 10-20 system. The reference electrode was placed on right mastoid and ground on AFz. The EMG electrodes were set at extensor digitorum, extensor carpi radialis, flexor carpi ulnaris, and palmaris longus. The reference was placed at the bone of the wrist. The cue and the data acquisition were carried out on MATLAB2015a and Curry8.0 of two computers respectively. For synchronization, MATLAB called the parallel port to send a trigger signal and it was recorded by Curry8.0 on another computer, when the participant was prompted to execute an action. Signals were sampled at 1000 Hz and filtered with a passband of 0.1-100Hz with a notch filter at 50Hz to remove power-line artifacts by EEGLAB toolbox in MATLAB. The moving time window method [21] was applied for EMG signals to determine the starting point of the movement. According to this time point, EEG signals were segmented to obtain the effective movement-related EEG data before and after movement onset.

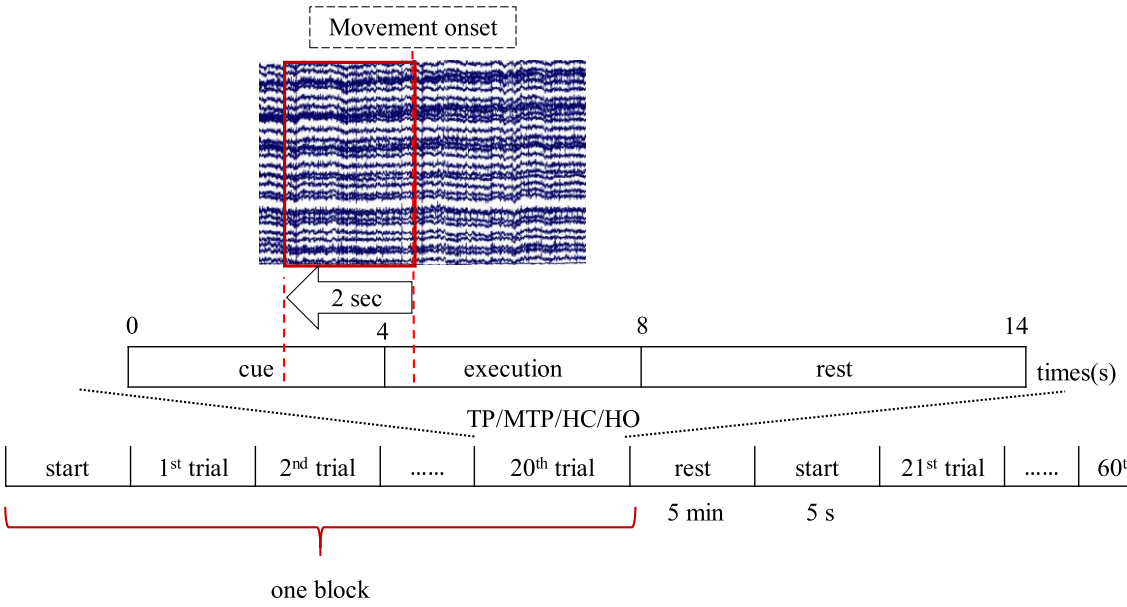

Fig. 1. Experimental protocol of one session. There were 4 sessions in total, corresponding to Tip Pinch, Multiple Tip Pinch, Hand Close and Hand Open four movements. One session consisted of three blocks, each of which comprised one start stage lasting 5 seconds and twenty trials lasting 14 seconds. A trial was consisted of a 4s-cue which indicated subject the movement they were going to perform, a 4s-execution which required subject only execute the specific movement and a 6s-rest which allowed subject to relax, blink, and swallow. There was a 5-min rest existed between the continuous blocks to prevent fatigue of the subject. The movement onset was finally determined by EMG using the sliding window method.

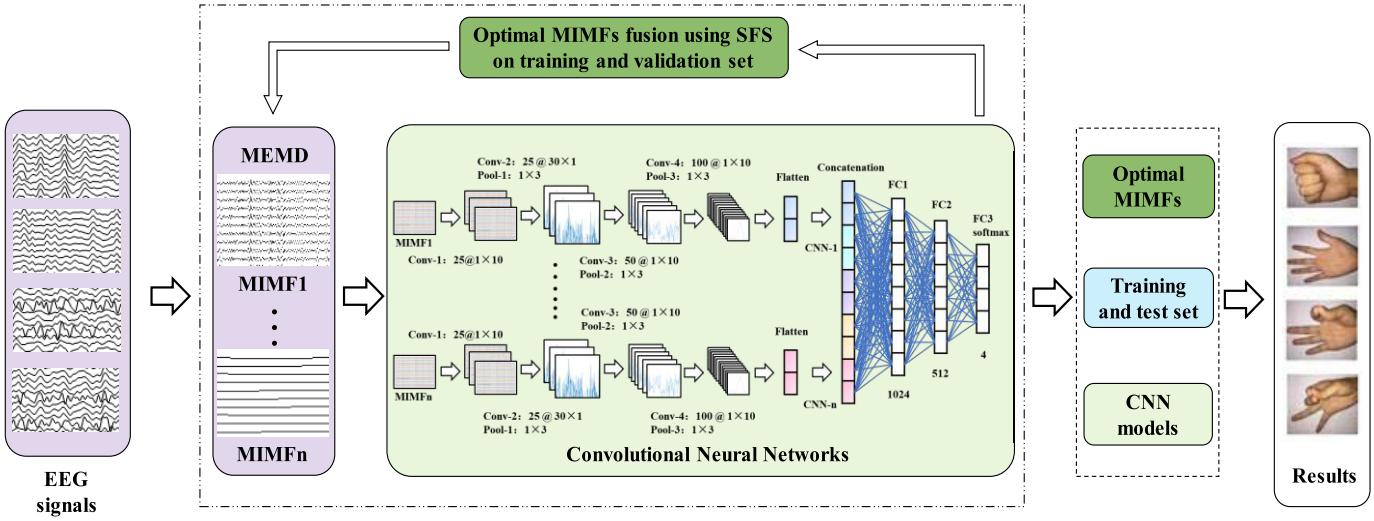

Fig. 2. Flowchart of the proposed algorithm.

B. Multi-Class EEG Recognition Algorithm

The flowchart of the proposed algorithm is shown in Fig.2. Raw EEG signals were divided into training set, validation set and test set. First, EEG signals were decomposed by MEMD into multiple multivariate intrinsic mode functions (MIMFs). Second, one MIMF was input to one CNN, and the CNN features would be concatenated after flattened [44]. The optimal MIMFs fusion was selected by SFS strategy on training and validation set. Finally, the selected MIMFs were input to the CNN for discriminating four kinds of hand movements. The average classification accuracy of six-fold cross-validation on test set were regarded as the performance evaluation of the algorithm.

1) Multivariate Empirical Mode Decomposition (MEMD): MEMD [45] is proposed based on empirical mode decomposition (EMD) [46] which overcomes the single-channel limitation of EMD by mapping low-dimensional multi-channel signals to higher dimensions. The MIMF component obtained by MEMD is an oscillating function with time-varying frequency, which reflect the local characteristics of non-stationary signals. The algorithm is essentially a smoothing process for non-stationary signals, and the n-dimension signals, $\{X(t)\}_{t=1}^{T} = \{x_1(t), x_2(t), \ldots, x_n(t)\}$, where T represents the length of the signals and t represents the time points of the signal, were gradually decomposed into several MIMFs, $\{h_i(t)\}_{i=1}^{q}$, and

TABLE I
MECN ARCHITECTURE

Layer	Filters	Size	Activation	Output Shape	Parameters
Input				(30,1000,1)	0
Conv2D	25	(1,10)	Linear	(30,991,25)	275
Conv2D	25	(30,1)	Linear	(1,991,25)	18775
BatchNorm				(1,991,25)	100
Actication			ELU	(1,991,25)	0
MaxPool2D		(1,3)		(1,330,25)	0
Conv2D	50	(1,10)	Linear	(1,321,50)	12550
BatchNorm				(1,321,50)	200
Actication			ELU	(1,321,50)	0
MaxPool2D		(1,3)		(1,107,50)	0
Conv2D	100	(1,10)	Linear	(1,98,100)	50100
BatchNorm				(1,98,100)	400
Actication				(1,98,100)	0
MaxPool2D		(1,3)		(1,32,100)	0
Flatten				(3200)	0
Dense	1024		Relu	(1024)	3277824
Dropout				(1024)	0
Dense	512		Relu	(512)	524800
Dropout				(512)	0
Dense	4		Softmax	(4)	2052

residue, $r\,(t)$, as in

$$X\,(t) = \sum_{i=1}^{q} h_i\,(t) + r\,(t)\,, \tag{1}$$

where q represents the number of MIMF decomposed. The specific calculation process is shown in literature [47]. Since data of different trials would be decomposed into different numbers of MIMF, the first seven MIMFs of each trial were selected as the output of MEMD, which were named MIMF1, MIMF2, MIMF3, …, MIMF6 and MIMF7 from high frequency to low frequency.

2) CNN Model: Convolutional neural network (CNN) is a feed-forward network which is inspired by the visual cortex of the human brain. It usually includes convolutional layer, pooling layer, and fully connected layer [48]. CNN can extract features automatically and it has been proven that a shallow CNN with large filter size can learn specific temporal and spatial features, while the deep model extracts general features [42].

The specific network structure is presented as Table I. In the first convolutional layer, there were 25 2D-convolutional kernels with size of 1×10. The second convolutional layer contained 25 2D-convolutional kernels with size of 30×1. Note that there was no activation function between the two layers so that the two layers could be seen as a temporal convolution and a spatial filter to achieve implicit regularization and reduce the amount of calculation. Then, a pooling layer (maximum pooling) was added behind the two convolutional layers. The size of the pool was 1×3, and the step stride was 1×3. The number of 2D-convolutional kernels of the third, fourth and fifth layers were set to 50, 100 and 200 with an identical pooling layer behind respectively. Next, the features extracted were flattened into one-dimension and the flattened data were fused by concatenation layer. In the subsequent selection of the optimal MIMFs fusion, the number of MIMF selected would be the number of constructed CNN. The CNN contained 3 fully connected layer with nodes 1024, 512, and 4. Softmax function was used in the last fully connected layer as the activation function to output the results of 4-class movement classification, and the Rectified Linear Unit (Relu, $f(x) = \max(x, 0)$) was used as the activation function of other fully connected layers. Considering that over-fitting occurs due to small sample size, the batch normalization layers were added between the convolution layer and the activation function [49]. Additionally, the dropout probability of the first two fully connected layers was set at 0.5 [50]. The batch size was set to 64. The exponential linear unit function (ELU, $f(x) = x \; for$ x>0 $and \; f(x) = e^x - 1 \; for \; x \leq 0$) was selected as the activation function of the convolutional layer to fit nonlinear problems. The cross-entropy loss function was selected as the loss function. The adaptive moment estimation (Adam) was chosen to upgrade the parameters of CNN, and the learning rate was 0.001, Beta1 was 0.9, and Beta2 was 0.999 according to the literature [51]. The early stopping method was set in the algorithm. The criterion for determining the optimal model was such that when the accuracy of the validation set data did not increase within 20 epochs, the model at this epoch was saved as the optimal model.

3) Optimal MIMFs Fusion Based on Sequential Forward Selection: After the original data were decomposed by MEMD, seven MIMFs were obtained and input to CNN-1, CNN-2, …, CNN-7. Features obtained after convolution and pooling were concatenated. But the feature concatenation brought some problems, such as the decline of generalization ability, the increase of training time and so on. Therefore, the MIMF selection algorithm was applied to select the proper MIMF and determine the optimal MIMFs fusion. The selection process consists of four parts: generation procedure, evaluation function, stopping criterion and validation procedure [52]. Sequential forward selection (SFS), sequential backward selection (SBS), and bidirectional search are commonly used strategies in generation procedure [53] and SFS was determined for the optimal MIMF combination selection in this study. The evaluation function is particularly important during the search process and classification accuracy was chosen. The stopping criterion was such that when the accuracy of the MIMF combination in the validation set was the highest, the MIMFs fusion was regarded as the optimal combination and the search stopped. Otherwise, next MIMF (from MIMF1 to MIMF7) would be added to calculate the accuracy.

4) Performance Assessment: In order to improve the generalization ability of the model, the pre-processed data were first divided into training set, validation set, and test set randomly of rate 3:2:1. The training set data were used to train the CNN and the validation set data were used for determining the optimal MIMFs fusion and the optimal model. The algorithm was tested on the test set and the final accuracy was averaged over the six-fold cross-validation.

Since the amount of data were too small for deep learning (data with a size of 60 × 30 × 2000 [trials × channels × sampling points] of one movement per subject after pre-processed), data augmentation was performed then to obtain data with a size of 1260 × 30 × 1000. The data were cropped with a one-second window, and the cropped step was set to 50 ms. The original 2-second-long EEG signals were cropped into twenty-one 1-second-long EEG signals finally. Since the training set, validation set, and test set had been divided before cropping, the crops with overlapping information would not affect the results. Overall, this resulted in 5040 (4 × 21 × 60) crops as CNN input, each of size 30 × 1000 per subject.

The average classification accuracy over six-fold cross-validation was used to evaluate the performance of the proposed algorithm in python 3.5. The Wilcoxon signed-rank tests were employed to check for statistical significance of accuracies between different data or methods.

III. Results

As we all know, the sensorimotor cortex area of the brain is activated during ME, so it is possible to classify hand motions by [29], [40], and [54] analyzing the EEG signals during ME. However, in recent years, studies [55] have shown that the action intention-related signals before ME also contain task-related information. Early decoding of motor states can reduce latency in practical applications, so as to provide stroke patients with more natural and active rehabilitation. In order to verify that the algorithm proposed in this paper can not only decode the EEG signals of ME, but also has a comparable classification performance for the EEG signals in the stage of movement preparation, the data before and after the movement were analyzed in the following parts: determination of CNN structure (Section III.A), algorithm performance verification (Section III.B), and different channel results (Section IV. Discussion).

A. Determination of CNN Structure

Considering that the main factors of the CNN structure are the size of the convolution kernel and the number of convolution layer, the two parameters were set to different values to explore the optimal network structure [56]. The size of the convolution kernel was set to 1 × 5, 1 × 10, 1 × 20, while the number of convolutional layers was set to 3, 4, and 5. That is, nine kinds of network structures were considered in total. The results of CNN network structure selection are shown in Fig. 3. The accuracy referred to the average result of 13 subjects over the six-fold cross-validation. BM in Fig. 3(a) meant that the accuracy was calculated by 2s-data before movement onset and AM in Fig. 3(b) was related to 2s-data after movement onset. From this figure, it could be observed that when the number of convolution kernel was 4 and the convolution kernel was selected as 1 × 10 ("4-10" in abscissa, Fig. 3(a)), the average accuracy (75.21±9.79%) was the highest among the nine network structures using data before movement onset. The accuracy of "4-10" (77.47±8.66%) in Fig. 3(b) was only 0.8% lower

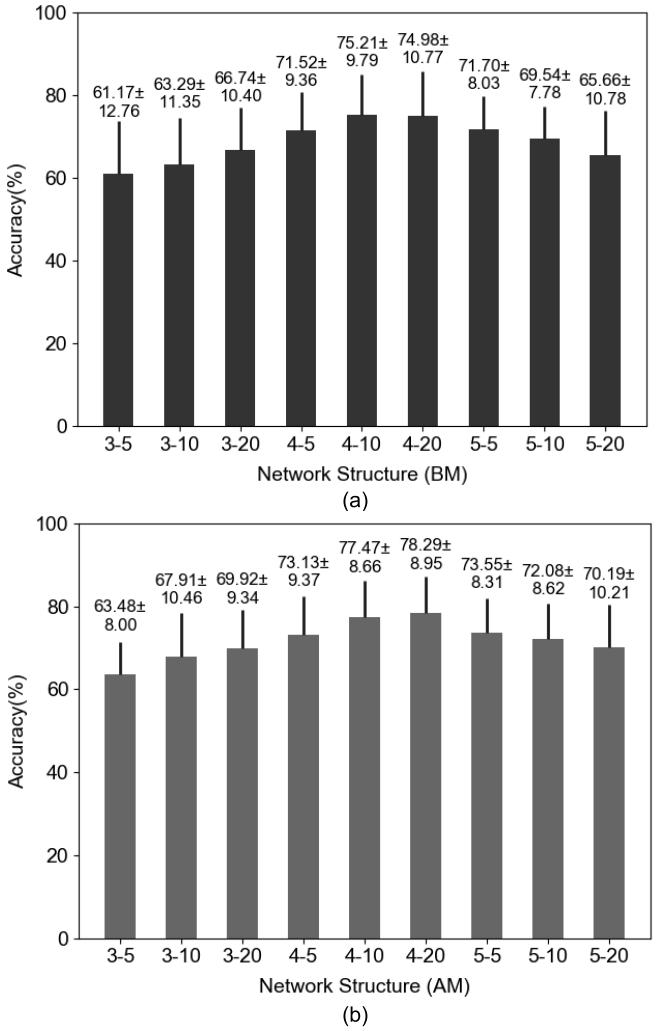

Fig. 3. CNN network structure selection before (a) and after (b) movement onset. The abscissa represents the network structure. The number before "-" represents the number of convolution layers and the number after "-" indicates the convolution core size. For instance, "4-10" represents a CNN with four convolution layers and each convolution core size is 1 × 10. BM stands for before movement onset and AM stands for after movement onset.

than that of "4-20" (78.29±8.95%). Considering that the computation time increased as convolution ernel became large, the network structure was confirmed to be "4-10" in this study.

B. Classification Accuracy of Movements Using Data Before and After Movement Onset

The decoding results of multi-class hand movements for 13 subjects are shown in Table II. The classification accuracies of different EEG signals were compared, including 2s data before the movement onset (BM-2s), 1s data before the movement onset (BM-1s), 1s data after the movement onset (AM-1s), and 2s data after movement onset (AM-2s). It could be seen that the accuracy of the four-class hand movements of BM-2s achieved 81.14%±6.76%, and the accuracy of AM-2s reached 81.08%±7.83%, which was slightly lower than that of BM-2s but there was no significant difference between them. The accuracies were 73.64±8.62% and 73.85±8.78%

for BM-1s and AM-1s respectively. The highest classification accuracy for single subject exceeded 90% and the lowest one was more than 65% when using 2s-long data, while the subject with the best classification performance had accuracies of nearly 90% and the worst accuracies were more than 55% using 1s-long data.

IV. Discussion

This study systematically evaluated the performance of the proposed algorithm MECN and showed that the MECN was able to decode the EEG signals of multi-class hand movements. The novelty and advantage of the proposed method are as follows: 1) MEMD was used firstly to decompose the EEG signals into several MIMFs with different frequency bands. Instead of reconstructing the signals by 2-3 MIMFs within a specific frequency band according to the experience in previous studies, different MIMFs were regarded as features in this paper; 2) The sequential forward selection (SFS) strategy was used to select MIMFs with the best result and inputting them into parallel CNNs, which can adaptively select the data of the best frequency band; 3) The first convolution layer of the CNN model works on the time dimension, then the second convolution layer performs a spatial filtering for all electrodes. Hence, the model extracts temporal and spatial features. The results show that the algorithm achieved good performance.

The complexity of model could be further analyzed from model framework, model size, optimization process, and data complexity [57]. The MECN model had only 4 convolution layers, each with 25, 25, 50, 100 convolution kernels, respectively. Adam optimizer was used to train the model because of the efficient computing and less memory required [51]. What's more, compared with inputting EEG signals as an image, we represented the input as a 2D-array while rows indicate channel and columns indicate time step, which reduced the input dimensionality. The MECN algorithm was performed on GeForce RTX 2080 GPUs with 8 GB memory. The machines had Intel(R) Xeon(R) E5-1650 v4 CPUs @ 3.60 GHz with 6 cores and 64 GB RAM. A subject-dependent model was trained on each subject, and the average time to train the model for 13 subjects was 3000.3 seconds. The training time mainly depended on the number of the optimal MIMF combinations selected. By contrast, the test time is much shorter, which has the potential for decoding in real-time BCI applications.

As shown in Table II, the average accuracies of 2s data achieved 81% (BM-2s: 81.14%±6.76%, AM-2s: 81.08%±7.83%), while the average accuracies of 1s data reached 70% (BM-1s: 73.64%±8.62%, AM-1s: 73.85%±8.77%). In order to compare the data of the same length in different time periods, the accuracies of the data of windows [−2s −1s] (71.99%±10.08%) and [1s 2s] (72.47%±8.09%) were additionally calculated (0s is regarded as the starting point of the action, the negative sign means the time before movement onset and the positive indicates the time after movement onset). The results show that the performance of windows [−2s −1s] and [1s 2s] were not significantly different from that of BM-1s (window [−1s 0s] vs [−2s −1s], $p = 0.13$) and AM-1s (window [0s 1s] vs [1s 2s], $p = 0.07$). All results of 1s data were significantly inferior to that of the 2s data. These show that the data in window [−2s −1s], [−1s 0s], [0s 1s], and [1s 2s] contain movement-related EEG information, and there is no doubt that more information useful for classification is contained within the [−2s 0s] and [0s 2s] signals. In addition, comparing the results before and after movement onset, there was little difference between these two kinds of data with the same length of time, indicating that the EEG signals prior to the movement onset already contained information that could be specifically classified. This is meaningful for the development of a low-latency prosthetic control or hand rehabilitation system in future research.

TABLE II
Classification Accuracies of 1-Second and 2-Second EEG Signals Before and After Movement Onset

Subject	BM-2s	BM-1s	AM-1s	AM-2s
1	81.17%	72.36%	74.33%	82.26%
2	82.74%	79.58%	78.63%	83.43%
3	**67.42%**	60.24%	**56.73%**	**65.46%**
4	73.85%	**59.84%**	64.80%	70.46%
5	79.33%	74.21%	71.03%	82.42%
6	79.35%	72.44%	73.35%	84.74%
7	82.44%	76.21%	76.03%	84.50%
8	93.00%	**88.29%**	**89.94%**	**92.02%**
9	80.81%	69.68%	69.52%	80.16%
10	82.62%	76.53%	76.01%	85.18%
11	82.38%	72.14%	69.78%	70.77%
12	**94.07%**	87.86%	88.31%	91.13%
13	75.62%	67.96%	71.61%	81.55%
Mean	81.14%	73.64%	73.85%	81.08%
Std	6.76%	8.62%	8.78%	7.83%
P-value	0.0015		0.0015	

Modern neurophysiological research had shown that the movement of human limbs could lead to the enhancement of brain signals activity in the corresponding areas of the sensorimotor cortex, which are mainly located in the parietal region [58]. In the EEG signal acquisition stage, we collected 30-channel EEG data of the whole brain. To investigate the influence of EEG channels on movement classification, Fig. 4 shows the comparison results of different number of channels before and after the movement onset. Fifteen-channel data from the fronto-parietal area (F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4, CP3, CPZ, CP4, P3, PZ, P4), and nine-channel data from frontal area (FC3, FCZ, FC4, C3, CZ, C4, CP3, CPZ, CP4) were analyzed in this comparison task. The average accuracy of 15 channels (BM: 66.90±7.83%, p=0.0024; AM: 66.77±14.16%, p=0.0030) and 9 channels EEG signals (BM: 61.09±12.43%, p=0.0015; AM: 61.20±13.97%, p=0.0015), regardless BM or AM, were statistically lower than that of thirty channels EEG signals (BM: 81.14%±6.76%, AM: 81.08±7.83%). Moreover, the accuracy of 15 channels EEG signals was higher than that of 9 channels EEG signals (BM: p=0.0015; AM: p=0.0107). These results suggested that more EEG channels contained more information, and the information related to the execution of hand movements was not

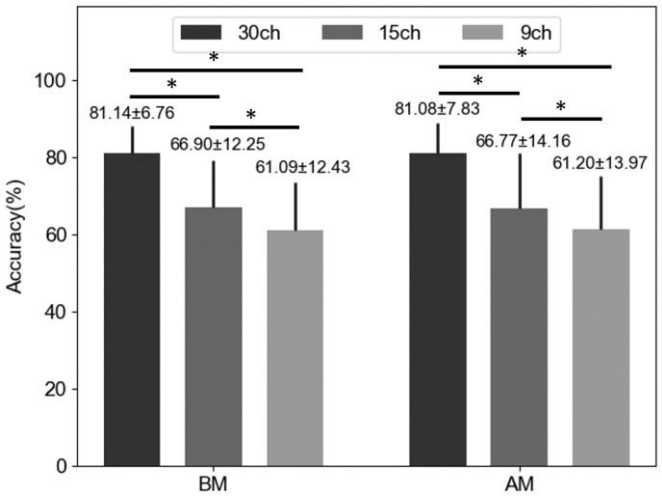

Fig. 4. Results of different channels before and after movement onset. Stars indicate statistically significant differences ($p<0.05$). Bars indicate standard deviation.

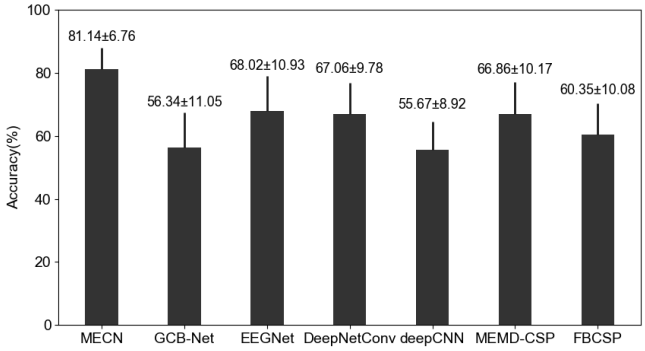

Fig. 5. Comparison of multi-class hand movement recognition using the proposed MECN method, GCB-Net, EEGNet, DeepNetConv, deepCNN, MEMD-CSP and FBCSP.

only comprised in the motor functional area. Fig. 3 shows the accuracy of inputting the raw data before and after the movement onset to different structure CNNs. The network structure of 4 convolutional layers and 1×10 convolution kernels were determined and the accuracy was 75.21±9.79%, and 77.47±8.66% before and after movement onset, respectively, which were significantly lower than that of data processed by the MECN algorithm (BM: $p=0.0071$; AM: $p=0.0277$). This indicated that MEMD and MIMFs fusion played an important role in classifying multi-class hand movements.

Fig. 5 presents the classification accuracy of the proposed method MECN and other advanced methods, such as GCB-Net [7], EEGNet [41], DeepNetConv [42], deepCNN [59], Filter Bank Common Spatial Pattern (FBCSP) [60], and MEMD-CSP [61]. These results were calculated using 2s-long data before movement onset. In the MEMD-CSP algorithm, the signals were first decomposed by MEMD. Second, the median frequency of MIMF was used to automatically find the subject specific MIMFs providing major contribution to mu and beta rhythms. Third, the selected MIMFs were summed to reconstruct the EEG signals. Fourth, CSP was used to extract the features of the reconstructed signals and the number of spatial filters was set at 5 ($m=5$). Finally, the features were classified by SVM classifier [62]. In the method of FBCSP, the signals were extracted features by CSP with filter band 0.1-10Hz, 10-20Hz, 20-30Hz, …, 80-90Hz, 90-100Hz, and classified by SVM. The average accuracies of the six compared methods were significantly lower than that of MECN ($p<0.05$). These results illustrated that the proposed method MECN was able to extract more suitable features and obtain higher classification accuracy than the state-of-the-art deep learning methods and traditional MEMD-CSP and FBCSP approach.

TABLE III
NUMBER OF TIMES EACH MIMF WAS SELECTED AS THE OPTIMAL MIMF COMBINATION IN THE SIX-FOLD CROSS-VALIDATION (EACH MIMF WAS SELECTED UP TO 6 TIMES EACH SUBJECT)

Subject \ MIMF	1	2	3	4	5	6	7
1	5	1	0	0	2	2	1
2	6	0	0	0	0	1	2
3	5	1	1	0	3	0	0
4	5	1	0	1	1	1	3
5	6	1	1	0	0	0	0
6	5	3	0	0	0	1	1
7	2	6	0	1	0	0	0
8	6	0	0	0	1	1	1
9	6	0	0	2	2	0	2
10	6	0	0	0	0	0	0
11	6	1	2	0	0	0	1
12	5	2	3	0	0	0	1
13	6	1	0	0	1	1	4
Total	69	17	7	4	10	7	16

The number of selected MIMFs was counted during optimal MIMFs fusion when using the 2s data before the movement onset. The number of times each subject's 7 MIMFs were selected are summarized in Table III. Since six-fold cross-validation was performed in the algorithm and one MIMF could be selected at most once in each fold cross-validation, each MIMF could be selected up to 6 times in total. It could be observed from this table that MIMF1 was selected the most times, followed by MIMF2. MIMF1 and MIMF2 of per subject were selected at least 5 times in the 6-fold cross-validation, which implied that MIMF1 and MIMF2 contained a lot of useful information which facilitated feature extraction and classification. However, MIMF7 had been selected 16 times, second only to the number of times MIMF2 had been selected. In order to explore the contribution of different MIMF to the results, single MIMF was input to CNN to get the corresponding classification accuracy. As shown in Fig. 6(a), MIMF1 reached the highest accuracy (79.4±7.05%) and MIMF7 reached the lowest accuracy (31.53±1.62%). Furthermore, as shown in Table III, except for S4 and S13, MIMF7 was selected only 1 or 2 times for other subjects. These indicated that there may be some information in the frequency band where MIMF7 was located on some subjects, but the information contained in MIMF7 was not very

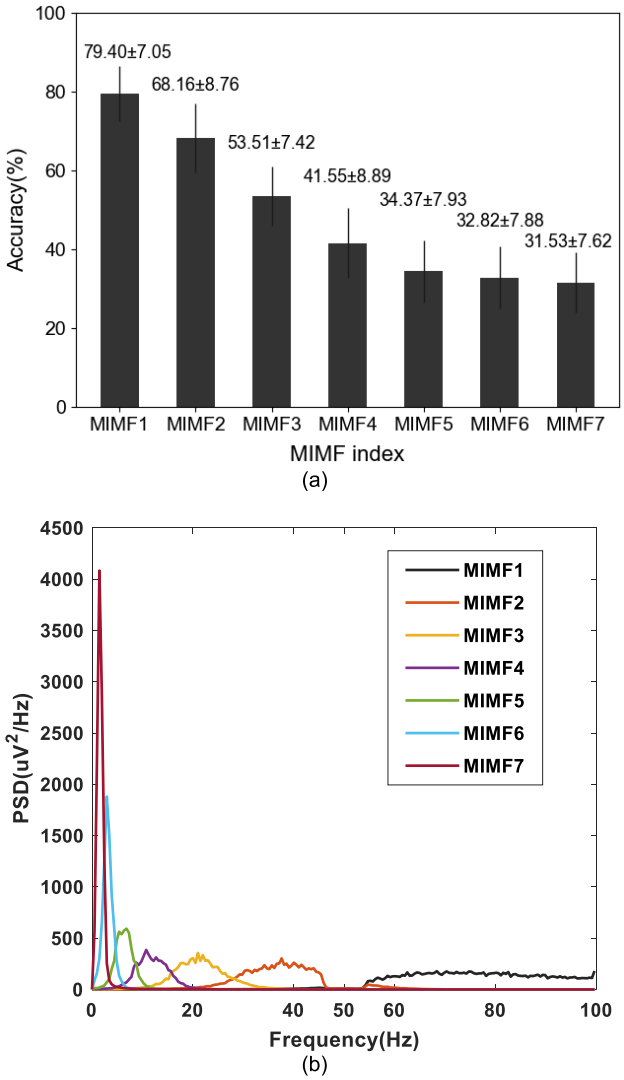

Fig. 6. (a) The accuracy of BM-2s using single MIMF with the algorithm proposed. Bars indicate standard deviation. (b) Average PSD of 60 trials by one subject (Hand Close) of seven MIMFs.

important for all subjects. Fig. 6(b) shows the power spectral density (PSD) of different MIMF frequency bands from one movement of one subject. MIMF1 contained high-gamma band information of 50-100Hz and MIMF2 contained low-gamma band information of about 30-45Hz. MIMF3, 4, 5, 6, 7 contained lower frequency band information than MIMF1 and MIMF2, while the classification accuracy got lower and lower. These demonstrated that the information of hand movements would be more contained in the gamma frequency band, which was consistent with previous literatures [63], [64], [65], [66], [67].

Furthermore, one limitation of the study was that the algorithm proposed in this paper was performed on healthy subjects, and this algorithm may not be equally applicable to stroke patients. In order to apply hand movement intention decoding to the rehabilitation training, we are going to collect the EEG signals from stroke patients and transfer learning would be considered to generalize the results to patient groups in the following study.

V. Conclusion

This study proposes a novel MECN algorithm to decode multi-class EEG signals of hand movement. The results showed that the accuracy obtained by the CNN with different layers and different convolution kernels were different. For the data collected in this experiment, the CNN with 4 convolution layers and 1×10 convolution kernel reached the best accuracy. The more the channels of EEG signals were, the more the useful information would be contained. The longer the length of EEG signals, the better the classification accuracy. The best accuracy of 81.14%±6.76% was achieved for 30 channels, 2s data before movement onset, which was significantly better than that of the traditional MEMD-CSP and FBCSP method. In addition, MIMF1 and MIMF2 were selected in a high percentage in MIMFs fusion, and they mainly referred to frequency band of 30-100Hz. This demonstrates that the movement-related EEG signals may be contained in the gamma band and the features of this band can be more considered in future research.

Acknowledgment

The authors acknowledge the support by the HPC platform of Xi'an Jiaotong University.

References

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophys.*, vol. 113, no. 6, pp. 767–791, 2002.

[2] A. Ramos-Murguialday *et al.*, "Brain–machine interface in chronic stroke rehabilitation: A controlled study," *Ann. Neurol.*, vol. 74, no. 1, pp. 8–100, Jul. 2013.

[3] Z. Yuan *et al.*, "Effect of BCI-controlled pedaling training system with multiple modalities of feedback on motor and cognitive function rehabilitation of early subacute stroke patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 2569–2577, 2021.

[4] M. Pritchard, A. I. Weinberg, J. A. R. Williams, F. Campelo, H. Goldingay, and D. R. Faria, "Dynamic fusion of electromyographic and electroencephalographic data towards use in robotic prosthesis control," in *Proc. J. Phys., Conf.*, Feb. 2021, vol. 1828, no. 1, Art. no. 012056.

[5] N. Yan *et al.*, "Quadcopter control system using a hybrid BCI based on off-line optimization and enhanced human-machine interaction," *IEEE Access*, vol. 8, pp. 1160–1172, 2020.

[6] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, no. 7753, pp. 493–498, Apr. 2019.

[7] T. Zhang, X. Wang, X. Xu, and C. P. Chen, "GCB-Net: Graph convolutional broad network and its application in emotion recognition," *IEEE Trans. Cybern.*, vol. 13, no. 1, pp. 379–388, Jan./Mar. 2022.

[8] M. A. L. Nicolelis and M. A. Lebedev, "Principles of neural ensemble physiology underlying the operation of brain–machine interfaces," *Nature Rev. Neurosci.*, vol. 10, pp. 530–540, Jul. 2009.

[9] A. Muralidharan, J. Chae, and D. M. Taylor, "Extracting attempted hand movements from EEGs in people with complete hand paralysis following stroke," *Frontiers Neurosci.*, vol. 5, p. 39, Mar. 2011.

[10] E. J. McDermott, J. Metsomaa, P. Belardinelli, M. Grosse-Wentrup, U. Ziemann, and C. Zrenner, "Predicting motor behavior: An efficient EEG signal processing pipeline to detect brain states with potential therapeutic relevance for VR-based neurorehabilitation," *Virtual Reality*, pp. 1–23, Sep. 2021, doi: 10.1007/s10055-021-00538-x.

[11] J. W. Krakauer, "Motor learning: Its relevance to stroke recovery and neurorehabilitation," *Current Opinion Neurol.*, vol. 19, no. 1, pp. 84–90, 2006.

[12] F. S. Miften, M. Diykh, S. Abdulla, S. Siuly, J. H. Green, and R. C. Deo, "A new framework for classification of multi-category hand grasps using EMG signals," *Artif. Intell. Med.*, vol. 112, Feb. 2021, Art. no. 102005.

[13] G. Jia, H.-K. Lam, J. Liao, and R. Wang, "Classification of electromyographic hand gesture signals using machine learning techniques," *Neurocomputing*, vol. 401, pp. 236–248, Aug. 2020.

[14] P. B. Shull, S. Jiang, Y. Zhu, and X. Zhu, "Hand gesture recognition and finger angle estimation via wrist-worn modified barometric pressure sensing," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 724–732, Apr. 2019.

[15] G. Wang and D. Ren, "Classification of surface electromyographic signals by means of multifractal singularity spectrum," *Med. Biol. Eng. Comput.*, vol. 51, no. 3, pp. 277–284, Mar. 2013.

[16] Y. Zhang, G. Wang, C. Teng, Z. Sun, and J. Wang, "The analysis of hand movement distinction based on relative frequency band energy method," *BioMed Res. Int.*, vol. 2014, pp. 1–8, Nov. 2014.

[17] G. Wang, Y. Zhang, and J. Wang, "The analysis of surface EMG signals with the wavelet-based correlation dimension method," *Comput. Math. Methods Med.*, vol. 2014, pp. 1–9, Apr. 2014.

[18] L. Cao, B. Xia, O. Maysam, J. Li, H. Xie, and N. Birbaumer, "A synchronous motor imagery based neural physiological paradigm for brain computer interface speller," *Frontiers Hum. Neurosci.*, vol. 11, p. 274, May 2017.

[19] F. Velasco-Álvarez, R. Ron-Angevin, L. D. Silva-Sauer, and S. Sancha-Ros, "Audio-cued motor imagery-based brain–computer interface: Navigation through virtual and real environments," *Neurocomputing*, vol. 121, pp. 89–98, Dec. 2013.

[20] R. Zhang *et al.*, "Control of a wheelchair in an indoor environment based on a brain–computer interface and automated navigation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 128–139, Jan. 2016.

[21] 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.

[22] E. Abdalsalam, M. M. Z. Yusoff, D. Mahmoud, A. S. Malik, and M. R. Bahloul, "Discrimination of four class simple limb motor imagery movements for brain–computer interface," *Biomed. Signal Process. Control*, vol. 44, pp. 181–190, Jul. 2018.

[23] D. Wang, D. Miao, and G. Blohm, "Multi-class motor imagery EEG decoding for brain–computer interfaces," *Frontiers Neurosci.*, vol. 6, p. 151, Oct. 2012.

[24] H. Yuan, T. Liu, R. Szarkowski, C. Rios, J. Ashe, and B. He, "Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: An EEG and fMRI study of motor imagery and movements," *NeuroImage*, vol. 49, no. 3, pp. 2596–2606, 2010.

[25] J. J. Daly and J. R. Wolpaw, "Brain–computer interfaces in neurological rehabilitation," *Lancet Neurology*, vol. 7, no. 11, pp. 1032–1043, Nov. 2008.

[26] Y. Suzuki *et al.*, "Muscle-specific movement-phase-dependent modulation of corticospinal excitability during upper-limb motor execution and motor imagery combined with virtual action observation," *Neurosci. Lett.*, vol. 755, Jun. 2021, Art. no. 135907.

[27] M. T. Carrillo-de-la-Peña, C. Lastra-Barreira, and S. Galdo-Álvarez, "Limb (hand vs. foot) and response conflict have similar effects on event-related potentials (ERPs) recorded during motor imagery and overt execution," *Eur. J. Neurosci.*, vol. 24, no. 2, pp. 635–643, Jul. 2006.

[28] B. Xu *et al.*, "Decoding hand movement types and kinematic information from electroencephalogram," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1744–1755, 2021.

[29] B. Gudiño-Mendoza, G. Sanchez-Ante, and J. M. Antelis, "Detecting the intention to move upper limbs from electroencephalographic brain signals," *Comput. Math. Methods Med.*, vol. 2016, pp. 1–11, Apr. 2016.

[30] H. Namazi, T. S. Ala, and V. Kulish, "Decoding of upper limb movement by fractal analysis of electroencephalogram (EEG) signal," *Fractals*, vol. 26, no. 5, Oct. 2018, Art. no. 1850081.

[31] C. Marquez-Chin, K. Atwell, and M. R. Popovic, "Prediction of specific hand movements using electroencephalographic signals," *J. Spinal Cord Med.*, vol. 40, no. 6, pp. 696–705, Nov. 2017.

[32] X. Lou, S. Xiao, Y. Qi, X. Hu, Y. Wang, and X. Zheng, "Corticomuscular coherence analysis on hand movement distinction for active rehabilitation," *Comput. Math. Methods Med.*, vol. 2013, pp. 1–10, Apr. 2013.

[33] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, "Decoding individual finger movements from one hand using human EEG signals," *PLoS One*, vol. 9, no. 1, Jan. 2014, Art. no. e85192.

[34] J. Wang, L. Bi, W. Fei, and C. Guan, "Decoding single-hand and both-hand movement directions from noninvasive neural signals," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 6, pp. 1932–1940, Jun. 2021.

[35] M. Jochumsen, I. K. Niazi, K. Dremstrup, and E. N. Kamavuako, "Detecting and classifying three different hand movement types through electroencephalography recordings for neurorehabilitation," *Med. Biol. Eng. Comput.*, vol. 54, no. 10, pp. 1491–1501, Oct. 2016.

[36] R. Alazrai, H. Alwanni, and M. I. Daoud, "EEG-based BCI system for decoding finger movements within the same hand," *Neurosci. Lett.*, vol. 698, pp. 113–120, Apr. 2019.

[37] S. H. Sardouie and M. B. Shamsollahi, "Selection of efficient features for discrimination of hand movements from MEG using a BCI competition IV data set," *Frontiers Neurosci.*, vol. 6, p. 42, Apr. 2012.

[38] F. Quandt, C. Reichert, H. Hinrichs, H. J. Heinze, R. T. Knight, and J. W. Rieger, "Single trial discrimination of individual finger movements on one hand: A combined MEG and EEG study," *NeuroImage*, vol. 59, no. 4, pp. 3316–3324, 2012.

[39] T. Chouhan, N. Robinson, A. P. Vinod, K. K. Ang, and C. Guan, "Wavlet phase-locking based binary classification of hand movement directions from EEG," *J. Neural Eng.*, vol. 15, no. 6, Sep. 2018, Art. no. 066008.

[40] N. Mammone, C. Ieracitano, and F. C. Morabito, "A deep CNN approach to decode motor preparation of upper limbs from time–frequency maps of EEG signals at source level," *Neural Netw.*, vol. 124, pp. 357–372, Apr. 2020.

[41] 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, Jul. 2018, Art. no. 056013.

[42] R. T. Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, pp. 5391–5420, Nov. 2017.

[43] J. Gutierrez-Martinez, J. A. Mercado-Gutierrez, B. E. Carvajal-Gámez, J. L. Rosas-Trigueros, and A. E. Contreras-Martinez, "Artificial intelligence algorithms in visual evoked potential-based brain-computer interfaces for motor rehabilitation applications: Systematic review and future directions," *Frontiers Hum. Neurosci.*, vol. 15, Nov. 2021, Art. no. PMC8656949.

[44] 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.

[45] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. Roy. Soc. A, Math., Phys., Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, Dec. 2009.

[46] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[47] Y. Tao, N. Yan, and W. Gang, "Hand movement prediction based on EEG signals by combining MEMD and CSP," in *Proc. 2nd Int. Conf. Image Process. Mach. Vis.*, 2020, pp. 105–112.

[48] 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.

[49] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[50] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

[51] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, 2015, pp. 1–15.

[52] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, nos. 1–4, pp. 131–156, 1997.

[53] X. Deng, Y. Li, J. Weng, and J. Zhang, "Feature selection for text classification: A review," *Multimedia Tools Appl.*, vol. 78, no. 3, pp. 3797–3816, 2019.

[54] A. Shakeel, M. S. Navid, M. N. Anwar, S. Mazhar, M. Jochumsen, and I. K. Niazi, "A review of techniques for detection of movement intention using movement-related cortical potentials," *Comput. Math. Methods Med.*, vol. 2015, pp. 1–13, Oct. 2015.

[55] A. Athanasiou, C. Lithari, K. Kalogianni, M. A. Klados, and P. D. Bamidis, "Source detection and functional connectivity of the sensorimotor cortex during actual and imaginary limb movement: A preliminary study on the implementation of eConnectome in motor imagery protocols," *Adv. Hum.-Comput. Interact.*, vol. 2012, pp. 1–10, Oct. 2012.

[56] T. Shi, L. Ren, and W. Cui, "Feature recognition of motor imaging EEG signals based on deep learning," *Pers. Ubiquitous Comput.*, vol. 23, nos. 3–4, pp. 499–510, Jul. 2019.

[57] X. Hu, L. Chu, J. Pei, W. Liu, and J. Bian, "Model complexity of deep learning: A survey," *Knowl. Inf. Syst.*, vol. 63, no. 10, pp. 2585–2619, Oct. 2021.

[58] N. Crone *et al.*, "Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band," *Brain*, vol. 121, no. 12, pp. 2271–2299, Dec. 1998.

[59] Z. Tang, C. Li, and S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks," *Optik*, vol. 130, pp. 11–18, Feb. 2017.

[60] 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.

[61] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "An automatic subject specific intrinsic mode function selection for enhancing two-class EEG-based motor imagery-brain computer interface," *IEEE Sensors J.*, vol. 19, no. 16, pp. 6938–6947, Aug. 2019.

[62] C. Chang and C. Lin, "LIBSVM," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, 2011.

[63] E. Demandt, C. Mehring, K. Vogt, A. Schulze-Bonhage, A. Aertsen, and T. Ball, "Reaching movement onset- and end-related characteristics of EEG spectral power modulations," *Frontiers Neurosci.*, vol. 6, p. 65, May 2012.

[64] M. Mohseni, V. Shalchyan, M. Jochumsen, and I. K. Niazi, "Upper limb complex movements decoding from pre-movement EEG signals using wavelet common spatial patterns," *Comput. Methods Programs Biomed.*, vol. 183, Jan. 2020, Art. no. 105076.

[65] G. Pfurtscheller, B. Graimann, J. E. Huggins, S. P. Levine, and L. A. Schuh, "Spatiotemporal patterns of beta desynchronization and gamma synchronization in corticographic data during self-paced movement," *Clin. Neurophysiol.*, vol. 114, no. 7, pp. 1226–1236, 2003.

[66] T. Ball *et al.*, "Movement related activity in the high gamma range of the human EEG," *NeuroImage*, vol. 41, no. 2, pp. 10–302, Jun. 2008.

[67] D. Popivanov, A. Mineva, and I. Krekule, "EEG patterns in theta and gamma frequency range and their probable relation to human voluntary movement organization," *Neurosci. Lett.*, vol. 267, pp. 5–8, May 1999.