IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 30, 2020, accepted August 16, 2020, date of publication August 24, 2020, date of current version September 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3018962

Diverse Feature Blend Based on Filter-Bank Common Spatial Pattern and Brain Functional Connectivity for Multiple Motor Imagery Detection

HONGTAO WANG[1], (Member, IEEE), TAO XU[1], CONG TANG[1], HONGWEI YUE[1], CHUANGQUAN CHEN[1], LINFENG XU[1], ZIAN PEI[1], (Graduate Student Member, IEEE), JIAJUN DONG[2], ANASTASIOS BEZERIANOS[3,4], (Senior Member, IEEE), AND JUNHUA LI[1,5], (Senior Member, IEEE)

[1]Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China
[2]Department of Neurosurgery, Jiangmen Central Hospital, Jiangmen 529030, China
[3]N.1 Institute for Health, National University of Singapore, Singapore 117456
[4]Department of Medical Physics, University of Patras, 26504 Patras, Greece
[5]School of Computer Science and Electronics Engineering, University of Essex, Colchester CO4 3SQ, U.K.

Corresponding authors: Hongtao Wang (nushongtaowang@qq.com) and Junhua Li (juhalee.bcmi@gmail.com)

This work was supported in part by the Science Foundation for Young Teachers of Wuyi University under Grant 2018td01, in part by the Jiangmen Brain-Like Computation and Hybrid Intelligence Research and Development Center under Grant [2018]359 and Grant [2019]26, in part by the Projects for International Scientific and Technological Cooperation under Grant 2018A050506084, in part by the Opening Foundation of Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging under Grant [2020]01, in part by the National Natural Science Foundation of China under Grant 61806149, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2020A1515010991, in part by the Startup Funds for Scientific Research of High-Level Talents of Wuyi University under Grant 5041700168, in part by the Innovation Training Foundation for Guangdong University Students under Grant 1081700308, and in part by the Innovation and Entrepreneurship Project of Wuyi University under Grant 3344100104.

ABSTRACT Motor imagery (MI) based brain-computer interface (BCI) is a research hotspot and has attracted lots of attention. Within this research topic, multiple MI classification is a challenge due to the difficulties caused by time-varying spatial features across different individuals. To deal with this challenge, we tried to fuse brain functional connectivity (BFC) and one-versus-the-rest filter-bank common spatial pattern (OVR-FBCSP) to improve the robustness of classification. The BFC features were extracted by phase locking value (PLV), representing the brain inter-regional interactions relevant to the MI, whilst the OVR-FBCSP is used to extract the spatial-frequency features related to the MI. These diverse features were then fed into a multi-kernel relevance vector machine (MK-RVM). The dataset with three motor imagery tasks (left hand MI, right hand MI, and feet MI) was used to assess the proposed method. Experimental results not only showed that the cascade structure of diverse feature fusion and MK-RVM achieved satisfactory classification performance (average accuracy: 83.81%, average kappa: 0.76), but also demonstrated that BFC plays a supplementary role in the MI classification. Moreover, the proposed method has a potential to be integrated into multiple MI online detection owing to the advantage of strong time-efficiency of RVM.

INDEX TERMS Multiple motor imagery, filter-bank common spatial pattern (FBCSP), phase locking value (PLV), brain functional connectivity (BFC), multi-kernel relevance vector machine (MK-RVM).

I. INTRODUCTION

Motor imagery (MI) is the imagination of actions and is associated with a specific activation in the brain. MI has been widely used in sport training, neurological rehabilitation, and brain-computer interface (BCI). For instance, the EEG-based MI BCI can enable a user to control a system based on the user's imagery movements of limbs [1]. Moreover, the MI BCI can be used in the stroke rehabilitation training [2]. Due to the high temporal resolution and non-invasive recording manner, EEG is widely used in brain studies and BCI applications. The brain activity recorded via EEG can be

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Zia Ur Rahman.

classified depending on the frequency of the signal. In particular, the alpha activity (8-12 Hz) and the beta activity (12-30 Hz) are mostly related to MI [3]. In addition, the increase or decrease in the activity at a certain frequency band locked to the onset of the MI refer to the event-related synchronization and desynchronization (ERS/ERD), respectively [4]. And thus, ERS and ERD are always used for characterizing MI [4]. Typically, the imaginary movements are four categories, which are left hand movement, right hand movement, movement of the feet, and movement of the tongue. More than two categories included in the classification forms multi-class classification. To extract effective features is quite important for the MI classification.

The feature extraction of MI-based EEG is a process that extracts discriminative information from filtered EEG signals. And such extracted information directly influences the classification accuracy of a classifier for MI. Typically, the basic feature extraction methods include the time-domain approaches and frequency-domain techniques. For example, autoregression and quaternions are typical time-domain methods and Fourier transform is based on frequency-domain [5]–[8]. However, independent use of time-domain or frequency-domain method may cause absence of features due to the ignorance of the spectral information or temporal features. Therefore, time-frequency domain techniques seem to satisfy both sides which combines spectral and temporal information together [9]. For example, Zhou *et al.* used a cascade structure in which the discrete wavelet transforms (DWT) decomposed EEG signals whereas the Hilbert transform (HT) could transform the decomposed one to wavelet envelop [10]. In addition to time-frequency domain techniques for feature extraction, the spatial domain analysis is also widely used for the classification of MI-based EEG signals [11]. Specifically, common spatial pattern (CSP) as well as its variants such as the common spatio-spectral pattern (CSSP) [12], the common sparse spatio-spectral patterns (CSSSP) and sub-band common spatial pattern (SBCSP) have attracted many interests [13], [14]. The variants of CSP aim to speed up the computational efficiency and improve classification accuracy. In this study, we propose a one versus the rest filter-bank common spatial pattern (OVR-FBCSP) method for the feature extraction of MI-based EEG.

The feature extraction can be considered in terms of the different domains of the signal. However, the information between different nodes of the electrode can show the property of neuron populations and thus neural connectivity should be paid attention during feature extraction. Recently, through the analysis of neural connectivity in the brain, the general function and communication between different regions of the brain are described. For example, Liang *et al.* [15] and Lee *et al.* [16] proved the functional connectivity in the process of motion imagination planning. Gong *et al.* proposed a brain network modeling method based on time-frequency cross mutual information of four classes of MI. Through statistical analysis and topological feature analysis, they observed significant differences in response level, response time, and activation target of four classes of MI tasks [17]. Moreover, Xu *et al.* proposed to use phase synchronization information to extract features to classify more than one category of MI of the same limb [18]. More remarkable is that Li *et al.* combined ERD/ERS analysis with dynamic networks in different MI stages to explore the dynamic processing of MI information. The results showed that the specific dynamic network structure conformed to the ERD/ERS evolution model [19].

Apart from the feature extraction method influences the accuracy and running speed of MI-based EEG classification, the classifier is another important issue. Typical classifier includes support vector machine (SVM), linear discriminant analysis (LDA), logistic regression, and artificial neural network (ANN) [8], [20], [21]. To further improve the classification accuracy of MI-based EEG signal, deep learning approach and recurrent neural network have been widely paid attention to [22], [23]. For example, Cheng *et al.* used deep neural network (DNN) as the classifier for the exploration of MI-based EEG pattern in stroke patients [24]. And DNN method (74.9%) gained a higher classification accuracy than that with SVM (67.7%). To achieve a trade-off between computational efficiency and classification accuracy, the multi-kernel method has attracted many interests [25]. In MI-based EEG classification, due to the variability of the EEG signal, single kernel function cannot be suitable for all imaginary movements. Here, we use a multi-kernel relevant vector machine (MK-RVM) to classify features of selected band subsets.

In this paper, we propose a cascade structure of one-versus-the-rest filter-bank common spatial pattern (OVR-FBCSP) method and MK-RVM for the classification of three imagery movements (left-hand movements, right-hand movements, and feet movements). This dataset was shared by the intelligent information processing and human-computer interaction laboratory at the Anhui University. The arrangement of this work is organized as follows. Section 2 shows the method of our study, which includes the experiment and data acquisition, signal preprocessing, feature extraction, feature selection, and classifier. Section 3 provides results of the method assessment. Discussions are presented in section 4. Finally, we give a conclusion in section 5.

II. METHOD

In this section, we focus on data acquisition and preprocessing, feature extraction, feature selection, and the classifier.

A. DATA ACQUISITION AND PREPROCESSING

The data were collected at the intelligent information processing and human-computer interaction laboratory of Anhui University. Six healthy subjects ranged from 22 to 28 years old participated in the MI experiment. During the experiment, subjects sit in front of a computer screen. The duration of each trial was 10 s and the onset of each trial was hinted with a ''beep'' tone.

Then the screen displays the cue arrows with the different directions which promote the subject to perform imagery movements (left-hand movements, right-hand movements, and feet movements). To prevent the subject from pre-imagining and to obtain stable EEG signals, the appearance of the arrows during the experiment was random with a duration of 6 s. After that, the computer displays a black screen and thus the subject can relax and wait for the next trial experiment. The data segments between 0.5-2.5 s (0.5-2.7 s for Subject 3, 0.5-2.6 s for Subject 5) relevant to the onset of visual cues presented to participants are used in this study. Each subject underwent 6 sessions, each of which comprises 75 trials (The numbers of trials for each category are not identical due to randomization). During the experiment, EEG signals were collected from the scalp using a headset with 14 electrodes (Fp1, Fp2, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, O1, Oz, and O2). Nine of them were used in this study (see Fig. 1). The sampling frequency was 250 Hz, and the data were sampled with a notch filter of 50 Hz and a band-pass filter of 0.5-100 Hz [26]. The details can be found in [27]. Then, we used the EEGLAB toolbox to do the data preprocessing, which included baseline subtraction, common average reference (CAR), and band-pass filtering. As the amplitude and spectrum power in μrhythm (8-12 Hz) and βrhythm (14-30 Hz) are mostly relevant to the MI, we chose six filter bands (7-12 Hz, 12-17 Hz, 17-22 Hz, 22-27 Hz, 27-32 Hz, 7-30 Hz) to extract features.

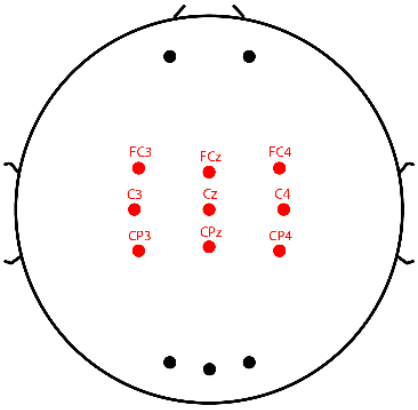


FIGURE 1. The layout of EEG electrodes according to the standard international 10-20 system. Red points are the electrodes used in this study.

B. FEATURE EXTRACTION

The feature extraction was performed by a combination of the FBCSP method and the brain functional connectivity (BFC) method.

1) FILTER-BANK COMMON SPATIAL PATTERN

The projection matrix $\mathbf{C}$ of CSP is constructed for filtering band data. In this paper, the first three pairs and the last three pairs of eigenvalues in the projection matrix are selected, and a total of six independent CSP filters are used for spatial filtering. Then we define the triple classes of MI EEG signals as $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$, respectively. Then the covariance model of multiple-class MI space is shown as following [28]:

$$\mathbf{R}_{\mathrm{i}} = \frac{\mathbf{X}_{\mathrm{i}}\mathbf{X}_{\mathrm{i}}^{\mathrm{T}}}{\mathbf{trace}(\mathbf{X}_{\mathrm{i}}\mathbf{X}_{\mathrm{i}}^{\mathrm{T}})}, \quad \mathbf{i} = 1, 2, 3 \tag{1}$$

The dimension of $\mathbf{X}_i$ is the multiplication of the number of channels, the time window, and the sampling frequency. $\mathbf{R}_i$ is the spatial covariance of EEG signals for each MI class. Then the composite covariance matrix can be denoted as:

$$\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3 \tag{2}$$

The singular value decomposition of covariance matrix $\mathbf{R}$ can be carried out as:

$$\mathbf{R} = \mathbf{U}_0\mathbf{\Lambda}_C\mathbf{U}_0^T \tag{3}$$

where $\mathbf{U}_0$ is the unitary matrix of principal components, and $\mathbf{\Lambda}_C$ is the diagonal matrix of eigenvalues. After calculating singular value decomposition of eigenvector and eigenvalue matrix, the transformation matrix of covariance matrix can be obtained:

$$\mathbf{Q} = \mathbf{\Lambda}_C^{-1/2}\mathbf{U}_0^T \tag{4}$$

In this paper, we need to calculate three OVR-CSP patterns. For example, the left-hand MI CSP versus the rest CSP can be calculated as:

$$\tilde{\mathbf{R}}_1 = \tilde{\mathbf{R}}_2 + \tilde{\mathbf{R}}_3 \tag{5}$$

Then we transform $\mathbf{R}_1$ and $\tilde{\mathbf{R}}_1$ into:

$$\mathbf{E}_1 = \mathbf{Q}\mathbf{R}_1\mathbf{Q}^T, \quad \tilde{\mathbf{E}}_1 = \mathbf{Q}\tilde{\mathbf{R}}_1\mathbf{Q}^T \tag{6}$$

Furthermore, we do eigenvalue decomposition for $\mathbf{E}_1$ and $\tilde{\mathbf{E}}_1$:

$$\mathbf{E}_1 = \mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^T, \tilde{\mathbf{E}}_1 = \mathbf{U}_1\tilde{\Lambda}_1\mathbf{U}_1^T \tag{7}$$

By combining the equation (1)-(7), we can get:

$$(\mathbf{Q}^T\mathbf{U}_1)^T\mathbf{R}_1(\mathbf{Q}^T\mathbf{U}_1) + (\mathbf{Q}^T\mathbf{U}_1)^T\tilde{\mathbf{R}}_1(\mathbf{Q}^T\mathbf{U}_1) = \mathbf{I} \tag{8}$$

where $\mathbf{U}_1$ is the common eigenvector matrix. Then we can get the CSP projection matrix $\mathbf{C}_1 = \mathbf{U}_1^T\mathbf{Q}$ and the selected features of left-hand MI can be obtained as following [29]:

$$\mathbf{f}_1 = \log\left(\frac{\mathrm{diag}(\mathbf{C}_1\mathbf{X}_i\mathbf{X}_i^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}})}{n}\right) \tag{9}$$

where n is the number of samples in the class i. Six frequency bands for each trial need 6×6 CSP filters for feature extraction. And six eigenvalues extracted from each frequency band are merged to obtain 36 eigenvalues. Then we carried out such feature extraction with normalization for each frequency band and thus the normalized feature can be obtained as:

$$F_i = \left\{\mathbf{f}_{i,1}, \mathbf{f}_{i,2}, \mathbf{f}_{i,3}, \mathbf{f}_{i,4}, \mathbf{f}_{i,5}, \mathbf{f}_{i,6}\right\}, \quad i = 1, 2, 3 \tag{10}$$

Furthermore, three classes of OVR-FBCSP features can be obtained as:

$$\mathbf{F}_{fbcsp} = [\mathrm{F}_1, \mathrm{F}_2, \mathrm{F}_3] \tag{11}$$

$\mathbf{F}_{fbcsp}$ of each trail is used to form 108 eigenvalues.

2) FUNCTIONAL BRAIN NETWORK

In this study, we use the phase locking value (PLV) to calculate brain connectivity as it has good performance in functional brain network analysis [18], [30]. The expression of PLV is as follows:

$$\mathbf{PLV} = \frac{1}{N}\left|\sum_{k=1}^{N} exp\Big(j\big(\sigma_1(t,k) - \sigma_2(t,k)\big)\Big)\right| \tag{12}$$

where N is the total number of sample points in a trial, $\sigma_1(t,k) - \sigma_2(t,k), k = 1, 2, \ldots, N$ represents the phase difference of a pair of electrodes at time t. Due to the binary network will lose a lot of information in the process of binarization, we choose the weighted network as the feature of this study. PLV value is between [0, 1], in which 0 is no connection between channels whereas 1 is the perfect connection between channels [18]. We use four measures for the evaluation of PLV which are degree, clustering coefficient, average path length, and local efficiency. These measurement methods are obtained by graph theory analysis through the brain connection toolbox [31].

(1) Degree: The degree of a node represents the number of edges of a node, which can be calculated with the following formula:

$$\mathbf{D}_i = \sum_{j \in G} \mathbf{w}_{ij} \tag{13}$$

where G is the set of all nodes in the network and $\mathbf{w}_{ij}$ is an element of the weighted network matrix. In the weighted network, the higher degree of a node, the more important it is.

(2) Clustering coefficient: The clustering coefficient indicates the clustering degree of brain function network nodes, which can be calculated with the following formula:

$$\mathbf{C}_i = \frac{2R}{\mathbf{D}_i(\mathbf{D}_i - 1)} \tag{14}$$

where R represents the number of neighboring nodes that directly connected to the node i.

(3) Average shortest path length: The average shortest path length reflects the information transferability within the brain function network, which can be expressed as:

$$\mathbf{L} = \frac{1}{M(M-1)} \sum_{i \geq j} \mathbf{L}_{ij} \tag{15}$$

where $\mathbf{L}_{ij}$ is the shortest path length between node i and j, M is the number of nodes.

(4) Local efficiency: Local efficiency is used to measure the ability of local information transmission and processing which can be expressed as:

$$\mathbf{E}_{local}(G) = \frac{1}{M} \sum_{i \in M} E_{global}(G_i) \tag{16}$$

where M is the number of nodes. $E_{global}(G_i)$ is the global efficiency of node i.

$$\mathbf{E}_{global}(G) = \frac{1}{M(M-1)} \sum_{i \neq j \in M} \frac{1}{\mathbf{L}_{ij}} \tag{17}$$

The measures of the PLV are extracted as the features of functional brain network.

3) FEATURE FUSION

The feature dimension of clustering coefficient and local efficiency extracted in each frequency band of each trail are Channels $\times$ Trials, while the feature dimension of average node degree and average shortest path length extracted in each frequency band of each test are $1\times$ Trials.

To check the effect of average shortest path length, we set up two combined schemes (PLV1: clustering coefficient, local efficiency, and average node degree; PLV2: clustering coefficient, local efficiency, average node degree, and average shortest path length) for the feature fusion with FBCSP [31]. Finally, FBCSP features and PLV complex network measure features were combined:

$$\mathbf{F} = [\mathbf{F}_{\text{fbcsp}}, \mathbf{F}_{plv}] \tag{18}$$

C. FEATURE SELECTION

In this study, we use the F-score feature selection method as it is a simple but effective method for the selection of the discriminative power of each feature in a feature set [33]. The dataset is divided into positive and negative classes. For example, the left hand is defined as positive class n_+ whereas the right hand and feet are defined as negative class n_-. We use F-score to sort the features and select the top n feature sets in the ranking list. And the value of n varies from 1 to half of the total [34]. Given the train dataset $x_k, k = 1, 2, \ldots, m$. Then, the F-score of the f^{th} feature of the dataset is defined as [35]:

$$S_f = \frac{(\bar{x}_f^{(+)} - \bar{x}_f)^2 + (\bar{x}_f^{(-)} - \bar{x}_f)^2}{\frac{1}{n_+ - 1}\sum_{k=1}^{n_+}\left(x_{k,f}^{(+)} - \bar{x}_f^{(+)}\right)^2 + \frac{1}{n_- - 1}\sum_{k=1}^{n_-}\left(x_{k,f}^{(-)} - \bar{x}_f^{(-)}\right)^2} \tag{19}$$

where $\bar{x}_f$, $\bar{x}_f^{(+)}$ and $\bar{x}_f^{(-)}$ are the mean value of the f^{th} feature on the whole training dataset, the mean value on the positive dataset, and the mean value on the negative dataset, respectively. $\bar{x}_{k,f}^{(+)}$ is the eigenvalue of the f^{th} feature of the k^{th} positive class, and $\bar{x}_{k,f}^{(-)}$ is the eigenvalue of the f^{th} feature of the k^{th} negative class. The great value of S_f indicates strong discrimination of features within different classes.

D. CLASSIFIER

In this paper, we use MK-RVM to realize multi-class MI recognition. The log marginal likelihood $\ell(\mathbf{A}) = \log P(\mathbf{Y}|\mathbf{K}, \mathbf{A}) = \log \int P(\mathbf{Y}|\mathbf{K}, \mathbf{W})P(\mathbf{W}|\mathbf{A})d\mathbf{W}$ which can

further be expressed as [36]:

$$\ell(\mathbf{A}) = \sum_{c=1}^{C} -\frac{1}{2}\left[N\log 2\pi + \log|\boldsymbol{C}| + \boldsymbol{y}_c^T\boldsymbol{C}^{-1}\boldsymbol{y}_c\right] \quad (20)$$

where $\mathbf{C} = \mathbf{I} + \mathbf{K}^T\mathbf{A}^{-1}\mathbf{K}$, $\mathbf{K}$ is the kernel. $\mathbf{A}$ is the scales matrix. $\boldsymbol{y}_c$ is the c^{th} column of regressor target $\mathbf{Y} \in \Re^{C\times N}$. $\mathbf{W} \in \Re^{C\times N}$ is the regressor that expresses weight for a specific class. And the log marginal likelihood can be also decomposed into:

$$\ell(\mathbf{A}) = \ell(\mathbf{A}_{-i}) + \sum_{c=1}^{C}\frac{1}{2}\left[\log\gamma_i - \log(\gamma_i + s_i) + \frac{q_{ci}^2}{\gamma_i + s_i}\right] \quad (21)$$

where s_i is the sparsity factor and q_{ci} is the new multi-class quality factor. The sparse factor is the measure of overlap between a sample $\mathbf{k}_i$ and the ones already included in the model. The quality factor q_{ci} measures the quality of a sample that describes a specific class. α_i is a hyperparameter.

We can get stationary points with the derivation $\partial\ell(\mathbf{A})/\partial\alpha_i = 0$ and α_i can be expressed as:

$$\alpha_i = \frac{Cs_i^2}{\sum_{c=1}^{C} q_{ci}^2 - Cs_i}, \quad if \ \sum_{c=1}^{C} q_{ci}^2 > Cs_i$$
$$\alpha_i = \infty, \quad if \ \sum_{c=1}^{C} q_{ci}^2 \le Cs_i \quad (22)$$

So $\mathbf{A}$ can be updated with Eq.(22) whereas the regressor $\mathbf{W}$ can update as:

$$\mathbf{W} = (\mathbf{K}\mathbf{K}^T + \mathbf{A})^{-1}\mathbf{K}\mathbf{Y}^T \quad (23)$$

In addition, $\mathbf{Y}$ can be updated as:

$$\tilde{y}_{cn} \leftarrow \hat{\mathbf{w}}_c^T\mathbf{k}_n - \frac{\varepsilon_{p(u)}\left\{N_u(\hat{\mathbf{w}}_c^T\mathbf{k}_n - \hat{\mathbf{w}}_i^T\mathbf{k}_n, 1)\Phi_u^{n,i,c}\right\}}{\varepsilon_{p(u)}\left\{\Phi(u + \hat{\mathbf{w}}_i^T\mathbf{k}_n - \hat{\mathbf{w}}_c^T\mathbf{k}_n)\Phi_u^{n,i,c}\right\}}, \quad c \neq i$$
$$\tilde{y}_{cn} \leftarrow \hat{\mathbf{w}}_c^T\mathbf{k}_n - (\sum_{j\neq i}\tilde{y}_{jn} - \hat{\mathbf{w}}_j^T\mathbf{k}_n), \quad c = i \quad (24)$$

where $\tilde{y}_{cn}$ is a standardized noise model, $u \sim N(0, 1)$ and Φ the Gaussian cumulative distribution function. $\mathbf{k}_n$ is each row of kernel $\mathbf{K}$.

In this paper, we use the 5-fold cross-validation to do the classification [6]. The whole procedure can be described as follows:

Step 1 Preprocessed each session was randomly divided into five sets, of which four sets are training samples (80%) and the remaining set is testing samples (20%).

Step 2 FBCSP and PLV measures were calculated, which was followed with feature fusion.

Step 3 Fused features were selected by the F-score feature selection method.

Step 4 The selected features were fed into the MK-RVM classifier.

E. PERFORMANCE MEASURE

Some measurement performances are used in this paper. The classification accuracy can be expressed by the following formula [34], [37]:

$$Accuracy = \frac{(TP + TN)}{N_{total}}$$
$$N_{total} = TP + TN + FP + FN \quad (25)$$

where TP is true positive, TN is true negative, FP is false positive, FN is false negative. $(TN + TP)$ is the number of correctly classified samples. N_{total} represents the total number of test samples.

Precision and recall are two measures widely used in the field of statistical classification, which are used to evaluate the classification results [37].

$$Precision = \frac{TP}{(TP + FP)} \quad (26)$$
$$Recall = \frac{TP}{(TP + FN)} \quad (27)$$

As a performance indicator of BCI, kappa is often used to measure multiple classes of problems. It is considered more robust than the overall agreement (accuracy) because it needs to take into account the chances of the agreement occurring [34]. We can be expressed as [38]:

$$Kappa = \frac{p_0 - p_e}{1 - p_e} \quad (28)$$

where p_0 is the overall agreement of the test samples, which is equal to the accuracy. p_e is the chance agreement probability value of the test samples, which can be obtained by the following formula:

$$p_e = \frac{\sum_i a_i b_i}{N_{total} \times N_{total}} \quad i = 1, 2, 3 \quad (29)$$

where a_i and b_i represent the sum of i^{th} class real samples and i^{th} class predicted samples of the confusion matrix, respectively. N_{total} is the total number of test samples.

The receiver operating characteristic (ROC) is an index to evaluate the performance of classifiers. In ROC space, the abscissa of each point is the false positive rate (FPR) and the ordinate is the true positive rate (TPR), which describes the trade-off between TP and FP.

$$FPR = \frac{FP}{(FP + TN)} \quad (30)$$
$$TPR = \frac{TP}{(TP + FN)} \quad (31)$$

where FPR represents the probability that a negative case is misclassified into a positive one, TPR represents the probability of pairing positive examples.

III. RESULTS

In the performance assessment, the data of each subject were divided into five equal sets randomly, of which four sets were for training and the remaining set was for testing.

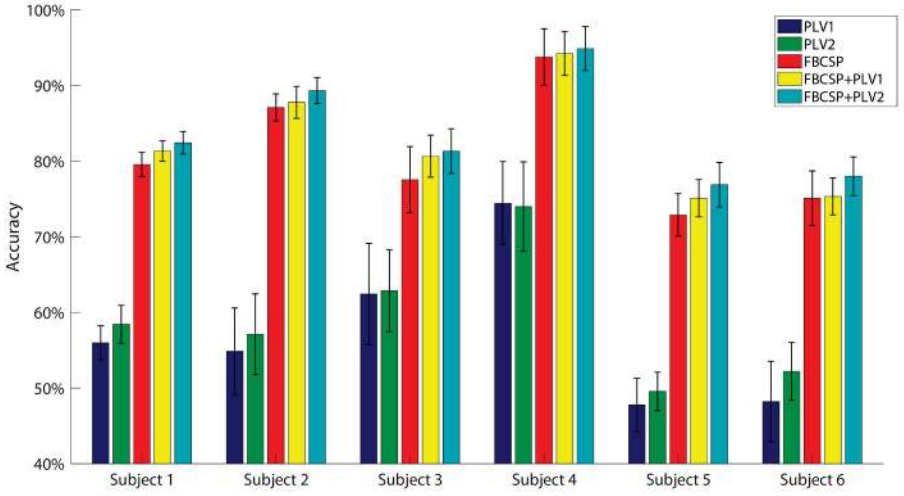


FIGURE 2. Classification accuracies for subject 1 to subject 6. The error-bar is the standard error of mean (SEM) of the accuracies of 5-fold cross-validation for each subject with different methods.

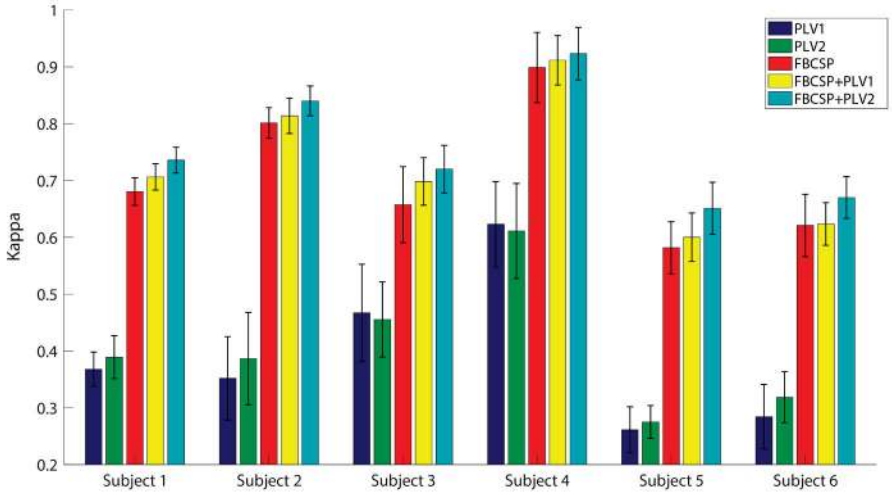


FIGURE 3. Kappa value for subject 1 to subject 6. The error-bar is the standard error of mean (SEM) of kappa of 5-fold cross-validation for each subject with different methods.

Each set was used as the testing samples one time to perform the 5-fold cross-validation. To demonstrate the classification performance, we compared the proposed method (FBCSP+PLV2) to the other four methods (PLV1, PLV2, FBCSP, FBCSP+PLV1, FBCSP+PLV2) and their accuracies are shown in Fig. 2. The best classification accuracy was found in Subject 4 and the classification accuracy of the proposed method is 94.89%±2.91% (Mean±SEM). Kappa as a measure multi-class classification performance [34] and the result is shown in Fig. 3. The kappa value for Subject 4 with the proposed method can reach 0.923±0.046.

To intuitively show the brain activity during MI, we draw topographical maps of OVR-FBCSP features extracted from EEG signals of subject 1. As mentioned in the section of feature extraction, selected features are transformed with corresponding projection matrix C in which the first three pairs of eigenvalues represent the selected features and the last three pairs are fused with the rest features. As can be seen from Fig. 4, when the subject does the leftward/rightward MI, the contralateral hemisphere is highly active. In addition, when the subject does the MI of feet, the central area of the selected nodes is active whereas the peripheral of the nine nodes is inactive with the combination of other two MI protocols (left-hand and right-hand movements). The connectivity matrix for each MI trial was constructed based on the EEG segment firstly, then they were divided to three groups according to the label. Fig. 5 shows mean connectivity matrices of left hand, right hand, and feet MI for Subject 1. The average connectivity matrix of left hand, right hand and feet MI is similar, and the diagonals of the three connectivity matrices are relatively large, which proves that the strong synchronization mainly occurs in temporal and parietal channels.

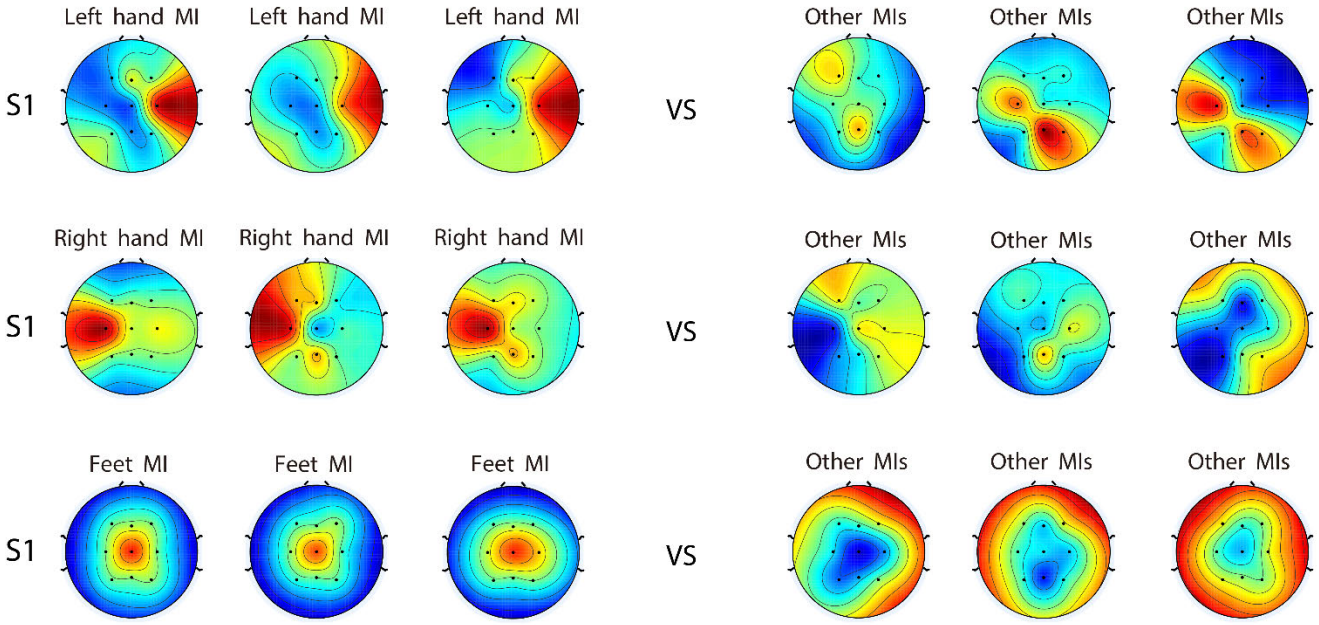


FIGURE 4. Topographies of FBCSP weights for subject 1. For each line, the left side is the FBCSP weights for MI of the left hand, right hand, and feet whereas the right side is FBCSP weights of the counterpart.

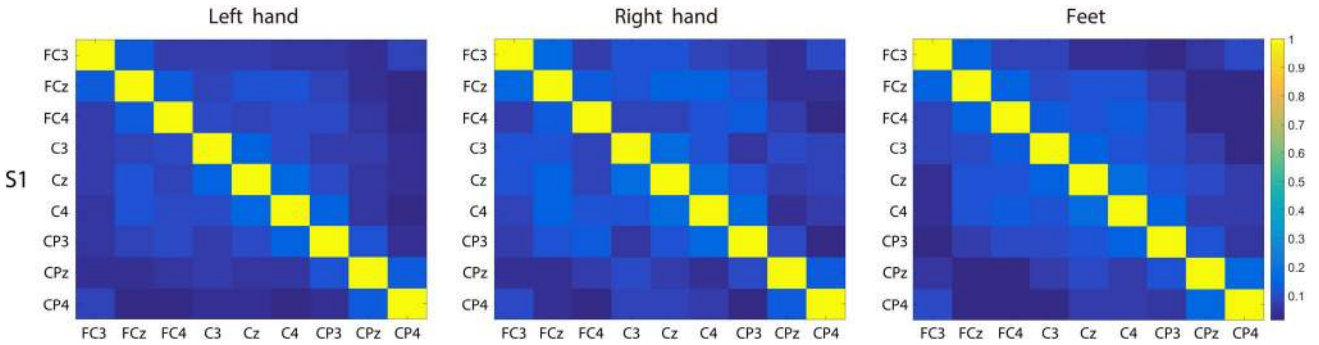


FIGURE 5. The mean connectivity matrix obtained from motor imagery EEG for Subject 1.

However, the upper right and lower left corners are dark blue, which indicates that the synchronization between frontal and occipital channels is weak.

Fig. 6 shows the confusion matrix of the sum of 5- fold cross-validation of 6 sessions for each subject under the FBCSP+PLV2 method. Fig. 7 shows the confusion matrix of the sum of 5-fold cross-validation of 6 sessions for each subject and the ROC curve of each subject and the average curve of three categories for each subject under the FBCSP+PLV2 method. We can see that this method has good classification performance and robustness. We calculate the precision and recall rate according to the confusion matrix under the FBCSP+PLV2 method, as shown in Table 1. Table 2 compares the FBCSP+PLV2 method with the other four methods and uses a t-test to calculate the p-value.

Furthermore, we also compared the time consumption among the adopted five methods (Table 3). The time cost is the averaged time of 5-fold cross-validation of the testing time for six subjects. The method combined by FBCSP and PLV2 has more features than the other four methods and thus takes a long-time cost. However, compared with SVM as the baseline, RVM showed stronger timely effectiveness.

TABLE 1. Averaged precision and recall of three categories for each subject (6-session for each subject) under the FBCSP+PLV2 method.

Subject	Precision (%)	Recall (%)
Subject 1	82.16	83.36
Subject 2	89.36	89.47
Subject 3	81.26	81.24
Subject 4	94.80	95.02
Subject 5	76.29	78.93
Subject 6	78.00	78.49

TABLE 2. Paired t-test (p-value) between the FBCSP+PLV2 and the other four methods.

T-test	PLV1	PLV2	FBCSP	FBCSP +PLV1	FBCSP +PLV2
p-value	1.16e-14	1.73e-15	0.1969	0.4829	-

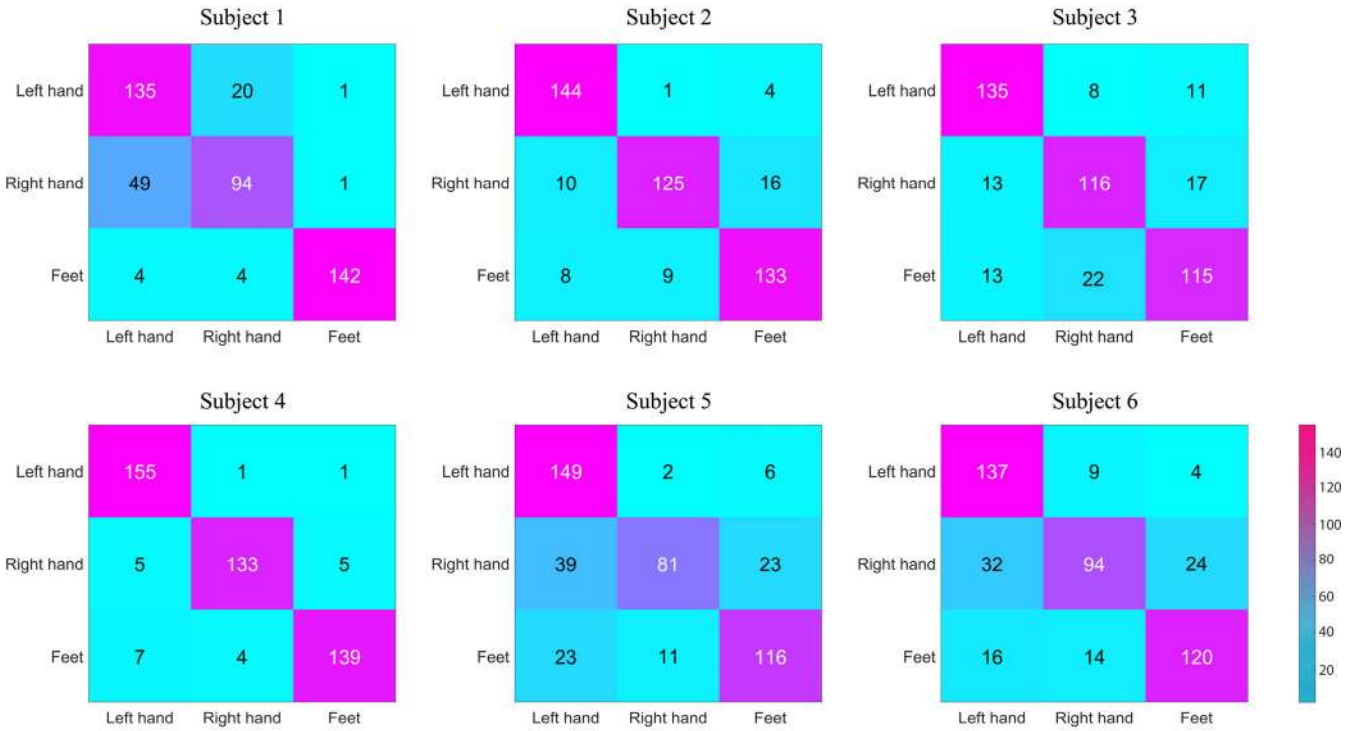


FIGURE 6. The confusion matrix obtained from motor imagery EEG for Subject 1 to Subject 6 under the FBCSP+PLV2 method. Each row represents the forecast category, and each column represents the real category.

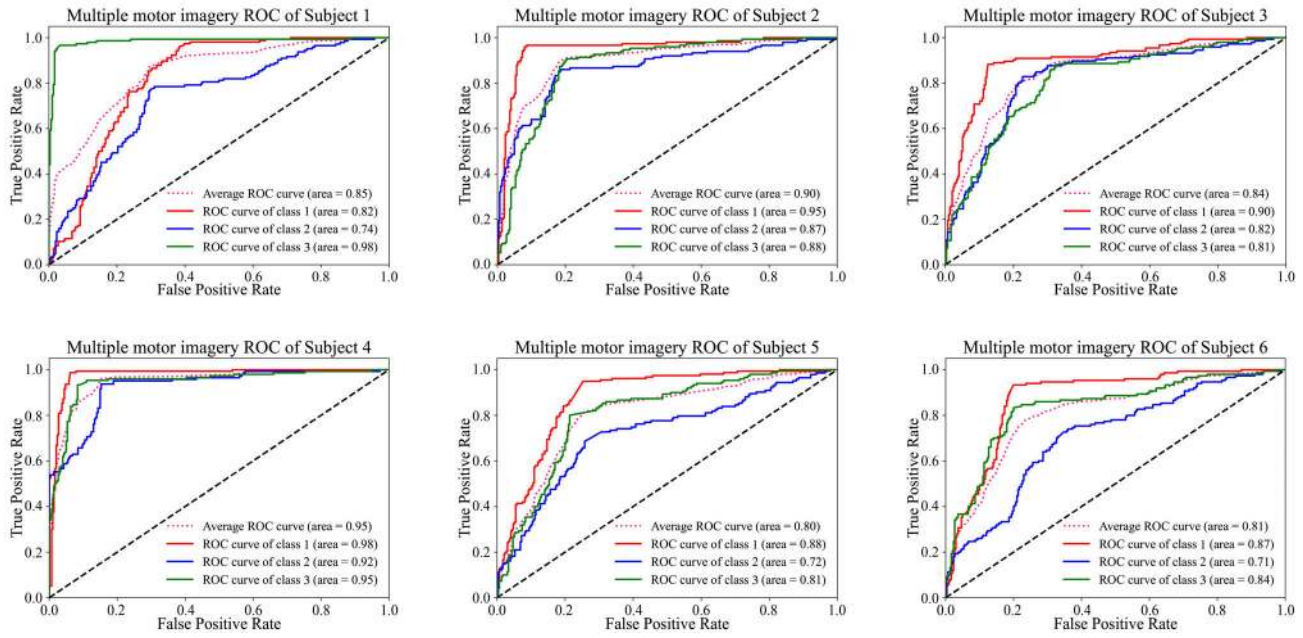


FIGURE 7. The multiple MI ROC obtained from motor imagery EEG for Subject 1 to Subject 6 under the FBCSP+PLV2 method, where Class1 is the left hand category, class2 is the right hand category, and class3 is the feet category.

IV. DISCUSSION

In this paper, we focus on classify three classes of MI data derived from the EEG signal. First, the collected data are pre-processed with the selection of the time window and frequency band. Then we extract features of pre-processed data, using the fusion of FBCSP and PLV complex network measure. Finally, we feed selected features into MK-RVM to classify the multiclass of MI. Here we aim to discuss the proposed classification structure in terms of time window optimization, feature fusion, classifiers, MI experiment, Comparison of different training protocols of MI tasks, and the limitation and prospect of the experiment.

TABLE 3. Averaged total testing time for each subject (6-session for each subject) with 5-fold cross-validation.

Methods	MK-RVM time cost (s)	SVM time cost(s)
PLV1	0.148	0.367
PLV2	0.161	0.383
FBCSP	0.145	0.359
FBCSP+PLV1	0.497	0.716
FBCSP+PLV2	0.515	0.752

A. TIME WINDOW SELECTION

The selection of the time window can remove the segmentation of EEG signals that have nothing to do with MI or eliminate data that is not the key time point of MI [27]. Our previous study has also shown that choosing the optimal time window for each subject can indeed improve the classification accuracy [39]. Although the optimal time windows of each subject's MI are different, it is found that the time between 0-2.5 s after the onset of visual cue can benefit the classification considering adequate sample points for subsequent data processing [40]. In this study, data between 0.5-2.5s (Subject 3 takes 0.5-2.7 s, Subject 5 takes 0.5-2.6 s) after the onset of the visual cue direction were used.

B. FEATURE FUSION STRATEGY

In this paper, by fusing the features of FBCSP and PLV complex network measure, we have achieved a good classification performance of multiple classes of MI. Feature fusion technique has also been applied in some other EEG-analysing related tasks. Ai *et al.* proposed a feature extraction method that combined the features of brain function network and local characteristic-scale decomposition (LCD) together. The good performance of this method was verified on the self-designed real-time BCI robot control and has put forward four classes of dataset [41]. In addition, to measure the complexity of EEG time series, Wang *et al.* proposed a fusion entropy (sample entropy, approximate entropy, and spectral entropy) analysis method for EEG and EOG signals. Results showed that the average accuracy of the fusion entropy analysis method combined with EOG and EEG can reach 99.1±1.2% [42]. Furthermore, Zhu *et al.* discussed the performance of multi-user MI-BCI idle detection based on common spatial pattern (CSP) and brain network features and proposed several cross-training feature fusion strategies [30]. The advantage of feature fusion was also found in other classification studies [43]–[45]. Based on the above studies, we can see that feature fusion outperforms single feature in the classification.

C. CLASSIFIER

The classifier selection plays a key role in the recognition of multiple MI tasks. MK-RVM has shown good performance in many aspects of classification tasks [36], [46], [47]. However, most of the studies on multi-class MI used multiple binary RVM. For example, Dong *et al.* added chaos dynamics into the kernel function of the RVM classifier in the framework of one versus one common spatial pattern (OVO-CSP) and thus made it excel in multi-class MI tasks [48]. Zhang *et al.* combined the location of EEG dipoles with CSP to extract

TABLE 4. Different method for MI classification.

Literatures	Dataset	Number of categories	Number of channels	Approach			Offline or real-time	Accuracy
				Feature extraction	Feature selection	Classifier		
She et al. [51]	▲	4	22	H-ELM	-	SS-ELM	Offline	67.76%
Ai et al. [41]	▲ □	4	CSP (22+ISC (9)) LCD (C3, Cz, C4) Brain network (22)	CSP+LCD+ Brain network	MCFS	SRDA	Offline Real-time	79.70% -
Razi et al.[38]	▲	4	21	CSP	MIBIF	DST-LDA	Offline	81.02%
Khan et al. [52]	○ ♦	3	8 14	SBCSP	SBFS	KNN NBPW	Offline	86.50% 60.61%
Zhou et al.[26]	□	3	Select the optimal channel for each subject	ICA	-	MRICs	Offline	76.60%
Dong et al. [50]	▲	4	22	PSCSP	-	RVM	Offline	74.39%
Xu et al. [53]	▲	4 5	22	ETR energy	-	CNN	Offline	85.57% 87.66%
The proposed	*	3	9	FBCSP+ PLV2	F-score	MK-RVM	Offline	84.27%

▲ BCI Competition IV Dataset 2a. □ Dataset acquired by themselves. ○ Wet gel electrodes. ♦ Emotiv Epoc. * Dataset provided by the intelligent information processing and human-computer interaction laboratory of Anhui University.
H-ELM: hierarchical extreme learning machine. SS-ELM: semi-supervised extreme learning machine. LCD: local characteristic-scale decomposition. MCFS: multicluster feature selection. SRDA: spectral regression discriminant analysis. MIBIF: mutual information-based best individual feature. DST-LDA: Dempster-Shafer theory linear discriminant analysis. SBCSP: sub-band common spatial patterns. SBFS: sequential backward floating selection. KNN: k-nearest neighbor. NBPW: Naïve-Bayesian Parzen-Window. ICA: independent component analysis. MRICs: motor-related independent components. PSCSP: phase space common spatial pattern. ETR: EEG topographical representation. CNN: convolutional neural network.

features from multi-class MI and extracted features were fed as the input to RVM [49]. Furthermore, Dong *et al.* proposed a new hybrid kernel RVM in which fused Gaussian kernel and polynomial kernel together. In MI tasks, by using the OVO-CSP strategy, phase space CSP (PSCSP) features were extracted and fed into the RVM classifier [50]. However, no relevant MK-RVM research has been found in multiple classes of MI tasks based on EEG signals and thus we use the MK-RVM classifier to classify three classes of MI tasks, and the best average accuracy can reach 83.81% (kappa: 0.76).

D. COMPARISON OF DIFFERENT TRAINING PROTOCOL OF MI TASKS

Here we summarize the related research of MI in recent years (Table 4). It can be seen from the Table 4 that the research on multiple classes of MI has gradually increased, but most of the research is still focused on offline data analysis, and there are few online analysis experiments. Compared with traditional classifiers and deep learning methods, the proposed method can achieve a medium to high classification accuracy.

We use FBCSP and PLV feature fusion, and then use MK-RVM for multiple MI recognition. Compared with the other latest methods (see Table 4), the method based on the feature fusion of FBCSP and PLV achieved better performance. This is due to the diverse features representing different MI-related information and help in the classification of MI tasks. The results demonstrated that the combination of FBCSP and PLV can extract discriminative features of both spatial-frequency and brain inter-regional interactions relevant to MI tasks. These complementary features could also benefit the understanding of underlying neural mechanisms of MI.

E. THE LIMITATION AND PROSPECT OF THE EXPERIMENT

There are some limitations to this study. First, our study performed the classification on three MI categories and the sample size for each category was not large. It would be better to evaluate the proposed method with more categories and more samples. Second, the proposed method is still an offline one and the processing of fused features takes a relatively long time. Moreover, the analysis has shown that in certain subjects, the selected channel number and frequency band have different optimal choices for different subjects [54]–[57]. Therefore, we will conduct more experiments on channel number and frequency bands in different subjects in future studies to verify this hypothesis.

V. CONCLUSION

In this paper, we proposed a cascade structure of feature fusion and MK-RVM for the classification of three-class MI tasks. The feature fusion integrates the OVR-FBCSP with PLV. The average classification accuracy reached 83.81%. The proposed method has a potential to be applied in real-time MI-based BCI applications.

ACKNOWLEDGMENT

(Hongtao Wang, Tao Xu, and Cong Tang contributed equally to this work.)

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] S. de Vries and T. Mulder, "Motor imagery and stroke rehabilitation: A critical discussion," *J. Rehabil. Med.*, vol. 39, no. 1, pp. 5–13, 2007.

[3] B. Graimann, B. Allison, and G. Pfurtscheller, "Brain–computer interfaces: A gentle introduction," in *Brain-Computer Interfaces*. Berlin, Germany: Springer, 2009, pp. 1–27.

[4] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. L. da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, May 2006.

[5] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomed. Signal Process. Control*, vol. 31, pp. 398–406, Jan. 2017.

[6] D. J. Krusienski, D. J. McFarland, and J. R. Wolpaw, "An evaluation of autoregressive spectral estimation model order for brain-computer interface applications," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 1323–1326.

[7] P. Batres-Mendoza, C. Montoro-Sanjose, E. Guerra-Hernandez, D. Almanza-Ojeda, H. Rostro-Gonzalez, R. Romero-Troncoso, and M. Ibarra-Manzano, "Quaternion-based signal analysis for motor imagery classification from electroencephalographic signals," *Sensors*, vol. 16, no. 3, p. 336, Mar. 2016.

[8] G. Rodríguez-Bermúdez and P. J. García-Laencina, "Automatic and adaptive classification of electroencephalographic signals for brain computer interfaces," *J. Med. Syst.*, vol. 36, no. S1, pp. 51–63, Nov. 2012.

[9] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, Feb. 2017, Art. no. 016003.

[10] J. Zhou, M. Meng, Y. Gao, Y. Ma, and Q. Zhang, "Classification of motor imagery EEG using wavelet envelope analysis and LSTM networks," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Jun. 2018, pp. 5600–5605.

[11] S. Kumar, A. Sharma, and T. Tsunoda, "An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information," *BMC Bioinf.*, vol. 18, no. S16, p. 545, Dec. 2017.

[12] S. Lemm, B. Blankertz, G. Curio, and K.-R. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 9, pp. 1541–1548, Sep. 2005.

[13] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Müller, "Combined optimization of spatial and temporal filters for improving brain-computer interfacing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2274–2281, Nov. 2006.

[14] Q. Novi, C. Guan, T. H. Dat, and P. Xue, "Sub-band common spatial pattern (SBCSP) for brain-computer interface," in *Proc. 3rd Int. IEEE/EMBS Conf. Neural Eng.*, May 2007, pp. 204–207.

[15] S. Liang, K.-S. Choi, J. Qin, Q. Wang, W.-M. Pang, and P.-A. Heng, "Discrimination of motor imagery tasks via information flow pattern of brain connectivity," *Technol. Health Care*, vol. 24, no. S2, pp. S795–S801, Jun. 2016.

[16] K.-B. Lee, K. K. Kim, J. Song, J. Ryu, Y. Kim, and C. Park, "Estimation of brain connectivity during motor imagery tasks using noise-assisted multivariate empirical mode decomposition," *J. Electr. Eng. Technol.*, vol. 11, no. 6, pp. 1812–1824, Nov. 2016.

[17] A. Gong, J. Liu, S. Chen, and Y. Fu, "Time–frequency cross mutual information analysis of the brain functional networks underlying multiclass motor imagery," *J. Motor Behav.*, vol. 50, no. 3, pp. 254–267, May 2018.

[18] 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.

[19] F. Li, W. Peng, Y. Jiang, L. Song, Y. Liao, C. Yi, L. Zhang, Y. Si, T. Zhang, F. Wang, R. Zhang, Y. Tian, Y. Zhang, D. Yao, and P. Xu, "The dynamic brain networks of motor imagery: Time-varying causality analysis of scalp EEG," *Int. J. Neural Syst.*, vol. 29, no. 1, Feb. 2019, Art. no. 1850016.

[20] M. Z. Baig, N. Aslam, H. P. H. Shum, and L. Zhang, "Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG," *Expert Syst. Appl.*, vol. 90, pp. 184–195, Dec. 2017.

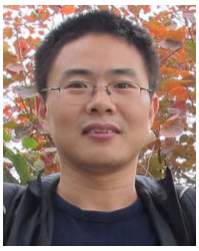
[21] Y. Zhang, Y. Wang, J. Jin, and X. Wang, "Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification," *Int. J. Neural Syst.*, vol. 27, no. 2, Mar. 2017, Art. no. 1650032.
[22] H. Yang, S. Sakhavi, K. K. Ang, and C. Guan, "On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2620–2623.
[23] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, Jun. 2017.
[24] D. Cheng, Y. Liu, and L. Zhang, "Exploring motor imagery EEG patterns for stroke patients with deep neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 2561–2565.
[25] Y. Zhang, Y. Wang, G. Zhou, J. Jin, B. Wang, X. Wang, and A. Cichocki, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," *Expert Syst. Appl.*, vol. 96, pp. 302–310, Apr. 2018.
[26] B. Zhou, X. Wu, J. Ruan, Z. Lv, and L. Zhang, "How many channels are suitable for independent component analysis in motor imagery brain-computer interface," *Biomed. Signal Process. Control*, vol. 50, pp. 103–120, Apr. 2019.
[27] B. Zhou, X. Wu, Z. Lv, L. Zhang, and X. Guo, "A fully automated trial selection method for optimization of motor imagery based brain-computer interface," *PLoS ONE*, vol. 11, no. 9, Sep. 2016, Art. no. e0162657.
[28] H. Wang, T. Li, A. Bezerianos, H. Huang, Y. He, and P. Chen, "The control of a virtual automatic car based on multiple patterns of motor imagery BCI," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 299–309, Jan. 2019.
[29] P. Bustios and J. L. Rosa, "Restricted exhaustive search for frequency band selection in motor imagery classification," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 4208–4213.
[30] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Hum. Brain Mapping*, vol. 8, no. 4, pp. 194–208, 1999.
[31] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.
[32] K. Keng 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.
[33] Q. Song, H. Jiang, and J. Liu, "Feature selection based on FDA and F-score for multi-class classification," *Expert Syst. Appl.*, vol. 81, pp. 22–27, Sep. 2017.
[34] S. Selim, M. M. Tantawi, H. A. Shedeed, and A. Badr, "A CSP/AM-BA-SVM approach for motor imagery BCI system," *IEEE Access*, vol. 6, pp. 49192–49208, 2018.
[35] W. Huang, H. Yan, R. Liu, L. Zhu, H. Zhang, and H. Chen, "F-score feature selection based Bayesian reconstruction of visual image from human brain activity," *Neurocomputing*, vol. 316, pp. 202–209, Nov. 2018.
[36] I. Psorakis, T. Damoulas, and M. A. Girolami, "Multiclass relevance vector machines: Sparsity and accuracy," *IEEE Trans. Neural Netw.*, vol. 21, no. 10, pp. 1588–1598, Oct. 2010.
[37] S. U. Kumar and H. H. Inbarani, "PSO-based feature selection and neighborhood rough set-based classification for BCI multiclass motor imagery task," *Neural Comput. Appl.*, vol. 28, no. 11, pp. 3239–3258, Nov. 2017.
[38] S. Razi, M. R. Karami Mollaei, and J. Ghasemi, "A novel method for classification of BCI multi-class motor imagery task based on Dempster–Shafer theory," *Inf. Sci.*, vol. 484, pp. 14–26, May 2019.
[39] H. Wang, C. Tang, T. Xu, T. Li, L. Xu, H. Yue, P. Chen, J. Li, and A. Bezerianos, "An approach of one-vs-rest filter bank common spatial pattern and spiking neural networks for multiple motor imagery decoding," *IEEE Access*, vol. 8, pp. 86850–86861, 2020.
[40] S. Kumar, K. Mamun, and A. Sharma, "CSP-TSM: Optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI," *Comput. Biol. Med.*, vol. 91, pp. 231–242, Dec. 2017.
[41] Q. Ai, A. Chen, K. Chen, Q. Liu, T. Zhou, S. Xin, and Z. Ji, "Feature extraction of four-class motor imagery EEG signals based on functional brain network," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026032.
[42] H. Wang, C. Wu, T. Li, Y. He, P. Chen, and A. Bezerianos, "Driving fatigue classification based on fusion entropy analysis combining EOG and EEG," *IEEE Access*, vol. 7, pp. 61975–61986, 2019.
[43] J. Li, Y. Wang, L. Zhang, A. Cichocki, and T.-P. Jung, "Decoding EEG in cognitive tasks with time-frequency and connectivity masks," *IEEE Trans. Cognit. Develop. Syst.*, vol. 8, no. 4, pp. 298–308, Dec. 2016.
[44] J. Li, Y. Wang, L. Zhang, and T.-P. Jung, "Combining ERPs and EEG spectral features for decoding intended movement direction," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 1769–1772.
[45] J. Harvy, E. Sigalas, N. Thakor, A. Bezerianos, and J. Li, "Performance improvement of driving fatigue identification based on power spectra and connectivity using feature level and decision level fusions," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 102–105.
[46] Y. Chen, T. Zhang, Z. Luo, and K. Sun, "A novel rolling bearing fault diagnosis and severity analysis method," *Appl. Sci.*, vol. 9, no. 11, p. 2356, Jun. 2019.
[47] W. He, "Recognition of human activities using a multiclass relevance vector machine," *Opt. Eng.*, vol. 51, no. 1, Feb. 2012, Art. no. 017202.
[48] E. Dong, G. Zhu, C. Chen, J. Tong, Y. Jiao, and S. Du, "Introducing chaos behavior to kernel relevance vector machine (RVM) for four-class EEG classification," *PLoS ONE*, vol. 13, no. 6, Jun. 2018, Art. no. e0198786.
[49] S. Zhang, R. Xu, A. N. Belkacem, D. Shin, K. Wang, Z. Wang, L. Yu, Z. Qiao, C. Wang, and C. Chen, "Classification of EEG multiple imagination tasks based on independent component analysis and relevant vector machines," in *IEEE MTT-S Int. Microw. Symp. Dig.*, vol. 1, May 2019, pp. 1–4.
[50] E. Dong, K. Zhou, J. Tong, and S. Du, "A novel hybrid kernel function relevance vector machine for multi-task motor imagery EEG classification," *Biomed. Signal Process. Control*, vol. 60, Jul. 2020, Art. no. 101991.
[51] Q. She, B. Hu, Z. Luo, T. Nguyen, and Y. Zhang, "A hierarchical semi-supervised extreme learning machine method for EEG recognition," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 147–157, Jan. 2019.
[52] J. Khan, M. H. Bhatti, U. G. Khan, and R. Iqbal, "Multiclass EEG motor-imagery classification with sub-band common spatial patterns," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, p. 174, Dec. 2019.
[53] M. Xu, J. Yao, Z. Zhang, R. Li, B. Yang, C. Li, J. Li, and J. Zhang, "Learning EEG topographical representation for classification via convolutional neural network," *Pattern Recognit.*, vol. 105, Sep. 2020, Art. no. 107390.
[54] H. Wang, T. Li, H. Huang, Y. He, and X. Liu, "A motor imagery analysis algorithm based on spatio-temporal-frequency joint selection and relevance vector machine," *Control theory Appl.*, vol. 34, no. 10, pp. 1403–1408, 2017.
[55] J. Li, C. Li, and A. Cichocki, "Canonical polyadic decomposition with auxiliary information for Brain–Computer interface," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 263–271, Jan. 2017.
[56] Y. Jiao, Y. Zhang, X. Chen, E. Yin, J. Jin, X. Wang, and A. Cichocki, "Sparse group representation model for motor imagery EEG classification," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 631–641, Mar. 2019.
[57] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Temporally constrained sparse group spatial patterns for motor imagery BCI," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3322–3332, Sep. 2019.

HONGTAO WANG (Member, IEEE) received the Ph.D. degree in pattern recognition and intelligent systems from the South China University of Technology in 2015.

He is currently a Full Professor and the Vice Dean of the Faculty of Intelligent Manufacturing, Wuyi University, and has been selected as a Thousand-Hundred-Ten Talent of Universities in Guangdong Province. He is also the Director of the Jiangmen Brain-Like Computation and Hybrid Intelligence Research and Development Center. From January 2017 to January 2019, he was a Visiting Research Fellow of the National University of Singapore. His current research interests include the fields of brain-like computation, pattern recognition, deep learning, and hybrid intelligence.

TAO XU received the Ph.D. degree from the Department of Biomedical Engineering, City University of Hong Kong, in 2019. He joined Wuyi University as a Distinguished Professor. His research interests include computational neuroscience and neural prosthetic systems.

CONG TANG was born in Hunan, China, in 1995. He received the bachelor's degree from the School of Xingxiang, Xiangtan University, China, in 2017. He is currently pursuing the master's degree with the Faculty of Intelligent Manufacturing, Wuyi University. His research interests include pattern recognition and brain-machine interface.



HONGWEI YUE received the Ph.D. degree in control theory and control engineering from the Guangdong University of Technology, China, in 2013. He is currently an Associate Professor with the Faculty of Intelligent Manufacturing, Wuyi University, China. His research interests include image processing, biomedical instruments, and information security.

CHUANGQUAN CHEN received the Ph.D. degree in computer science from the University of Macau, Macau, China. He subsequently joined Wuyi University as a Distinguished Professor. His current research interests include machine learning methods and intelligent systems.

LINFENG XU was born in Guandong, China, in 1994. He received the bachelor's degree from the Guangdong University of Petrochemical Technology, China, in 2017. He is currently pursuing the master's degree with the Faculty of Intelligent Manufacturing, Wuyi University. From 2017 to 2019, he was involved in algorithm research and development in Shenzhen. His research interests include pattern recognition and brain-like computation.

ZIAN PEI (Graduate Student Member, IEEE) was born in Hebei, China, in 1996. He received the bachelor's degree in building electricity and intelligence from the Hebei University of Architecture, China, in 2018. He is currently pursuing the master's degree with the Faculty of Intelligent Manufacturing, Wuyi University. He won the P300-based BCI competition at the recent BCI Controlled Robot Contest of World Robot Conference held in Beijing in 2019. His current research interests include the fields of brain-like computation, pattern recognition, and hybrid intelligence.

JIAJUN DONG received the Ph.D. degree in neurosurgery from the University of Soochow, Suzhou, China, in 2006. Later, he worked as a Postdoctoral Researcher with Sun Yat-sen University. He is currently the Chief Physician of neurosurgery with the Jiangmen Central Hospital. He has published more than 30 articles as the first and second author and participated in the compilation of two works. His current research interest includes solving the common, frequently occurring, and complicated clinical problems in neurosurgery.

ANASTASIOS BEZERIANOS (Senior Member, IEEE) received the degree in physics from the University of Patras, Patras, Greece, the degree in telecommunications from Athens University, and the Ph.D. degree in bioengineering from the University of Patras. He is currently a Professor with the N.1 Institute for Health, National University of Singapore, and a Visiting Professor with the Department of Computer Science, New South Wales University (NSWU), Canberra, ACT, Australia. He has been a Professor of medical physics with the Medical School, University of Patras, since 2004. His research interests include diverse areas spanning from artificial intelligence and robotics to biomedical signal processing, brain imaging, mathematical biology and systems medicine, and bioinformatics. His work was summarized in 140 journal articles and 217 conference proceedings publications and one book, and he holds two patents. He has research collaborations with research institutes and universities in Japan, China, and Europe. He is also a Fellow of the European Alliance for Medical and Biological Engineering and Science (EAMBES), and the Founder and the Chairman of the Biannual International Summer School on Emerging Technologies in Biomedicine. He is also an Associate Editor of the IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING (TNSRE), the *PLOS One*, and *Neuroscience* journal, and a Reviewer for several international scientific journals. He is also a Registered Expert of the Horizon 2020 Program of the European Union and a Reviewer of research grant proposals in Greece, Italy, Cyprus, and Canada.

JUNHUA LI (Senior Member, IEEE) received the Ph.D. degree in computer science from the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China, in 2013. He is currently a Lecturer with the School of Computer Science and Electronics Engineering, University of Essex, Colchester, U.K. He was a Senior Research Fellow of the National University of Singapore, Singapore. His research interests include computational neuroscience, brain-computer interface, machine learning, and neurophysiological data analytics and their practical applications. He is an Associate Editor of IEEE ACCESS and a Review Editor of *Frontiers in Human Neuroscience*, and served as a guest associate editor of several special issues related to his research interests.