

Research Article

Multiclass Motor Imagery Recognition of Single Joint in Upper Limb Based on NSGA- II OVO TWSVM

Shan Guan , Kai Zhao , and Fuwang Wang 

School of Mechanical Engineering, Northeast Electric Power University, 132012 Jilin, China

Correspondence should be addressed to Shan Guan; guanshan1970@163.com

Received 19 April 2018; Accepted 7 June 2018; Published 28 June 2018

Academic Editor: Plácido R. Pinheiro

Copyright © 2018 Shan Guan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the study of the brain computer interface (BCI) system, electroencephalogram (EEG) signals induced by different movements of the same joint are hard to distinguish. This paper proposes a novel scheme that combined amplitude-frequency (AF) information of intrinsic mode function (IMF) with common spatial pattern (CSP), namely, AF-CSP to extract motor imagery (MI) features, and to improve classification performance, the second generation nondominated sorting evolutionary algorithm (NSGA-II) is used to tune hyperparameters for linear and nonlinear kernel one versus one twin support vector machine (OVO TWSVM). This model is compared with least squares support vector machine (LS-SVM), back propagation (BP), extreme learning machine (ELM), particle swarm optimization support vector machine (PSO-SVM), and grid search OVO TWSVM (GS OVO TWSVM) on our dataset; the recognition accuracy increased by 5.92%, 22.44%, 22.65%, 8.69%, and 5.75%. The proposed method has helped to achieve higher accuracy in BCI systems.

1. Introduction

BCI is a technology that enables the brain to establish communication and control directly between human brain and computer or other electronic devices without the help of peripheral nerves and limbs [1, 2]. BCI technology not only enhances the ability of disabled patients to communicate with the outside world in the field of medical rehabilitation [3, 4], but also has wide applications in smart home, mass consumption and entertainment, military, and other fields. At present, the research direction of BCI system is mainly in the following aspects: sensorimotor (SMR) [5], slow cortical potential (SCP) [6], P300 event-related potential [7], and steady-state visual evoked potential (SSVEP) [8]. The most widely used is the SMR BCI system based on motor imagery.

The μ (8-13Hz) and β (13-30Hz) rhythms in EEG signals will cause a phenomenon named event-related desynchronization (ERD) and event-related synchronization (ERS) when motor imagery occurs [9, 10]. This means that the rhythmic activities of the brain represent frequency specific changes may consist either of decreases or of increases of

power in given frequency bands. The ERD/ERS phenomenon is an important basis for the BCI systems of motor imagery.

In order to improve the classification accuracy of the BCI system, researchers have studied the feature extraction methods and classification methods of EEG signals [11–15]. The most commonly used feature extraction methods include wavelet packet transform (WT), Fourier transform (FT), CSP [16, 17], and autoregressive (AR) model. The classification methods include linear discriminant analysis (LDA), support vector machine (SVM), neural network (NN), and so on. Wang et al. used the convolution neural network (CNN) to recognize the image of the brain topographic map of three kinds of motor imagery movements of the upper limb, flexion wrist, and wrist external rotation, and the highest recognition rate in the three classification experiments is 67.89% [18]. Roy et al. carried out Hilbert transform for two kinds of motor imagery of shoulder and elbow joint and used discrete wavelet transformation to extract features; SVM gets the highest recognition rate of 84.91% in the five recognition methods [19]. Sachin et al. used empirical mode decomposition (EMD) to extract the energy features of left and right hand motor

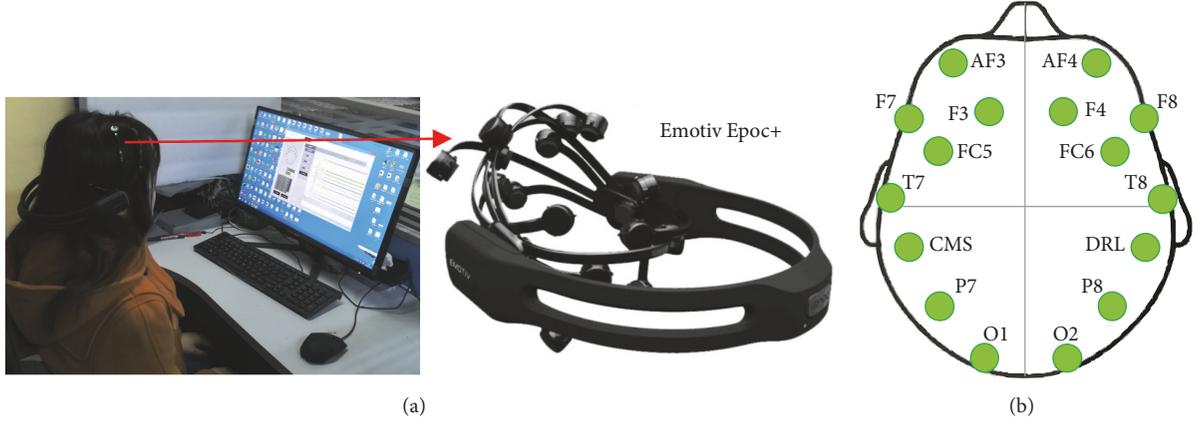


FIGURE 1: (a) Experimental photos and Emotiv Epoc+ and (b) Emotiv 14 electrodes located over 10-20 international system positions.

imagery EEG signals and classified them using LS-SVM; the recognition rate is no less than 95.56% [20]. Tang et al. used PSO to optimize the hidden-layer visible deep stacking network (HVDSN) to recognize the left and right hand motor imagery EEG signals; the recognition rate is no less than 89.84% [21]. Although the above studies have achieved high recognition results, most of the research focuses on hands and feet motor imagery, few studies have been conducted based on multiclass motor imagery of single joint.

In this paper, we propose a method to improve accuracy of motor imagery BCI using AF-CSP and an optimized OVO TWSVM classifier. The proposed method is composed of a total of four stages. First, a notch filter and common average reference (CAR) are used to remove noise in EEG signal. Second, EMD is used to obtain IMF, and FFT is used to obtain AF information of the IMF. Third, the NSGA-II is used to tune hyperparameters for linear and nonlinear OVO TWSVM. Finally, an optimized OVO TWSVM classifier is evaluated using laboratory data sets (three kinds of motor imagery of shoulder flexion, extension, and abduction) and compared with state-of-the-art algorithms (LS-SVM, BP, ELM, PSO-SVM, and GS OVO TWSVM).

2. Materials and Methods

Emotiv Epoc+ is used to collect EEG data of motor imagery. It is a portable EEG acquisition device with a sampling rate of 128Hz. It has fourteen electrode channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4) and two inference electrodes (CMS, DRL), and the electrode placement follows the international 10-20 standard. Experimental photos, equipment, and the Emotiv 14 electrodes are located over 10-20 international system positions as shown in Figure 1. This experiment collected three kinds of EEG signals of one joint: imagination of shoulder flexion, extension, and abduction, as shown in Figure 2.

Seven subjects participated in this experimental study. These subjects were in good health. During the experiment, subjects were naturally placed with both hands, trying to avoid body or head movement. During the experiment, subjects carried out motor imagery under the outside cue,

a single experiment collected EEG signal for 5 seconds, and then take 5-7 seconds to have a rest; each action repeated acquisition 20 times. The experimental process is shown in Figure 3. The dimension of EEG is high and the amount of data is large, in order to reduce the computational complexity. In this paper, the EEG signals collected from four electrode channels of FC5, F3, F4, and FC6 were selected for the following motor imagery analysis.

3. Theories and Methods

3.1. Data Preprocessing. EEG signals contain a variety of noise, and it is necessary to perform spatial filtering before feature extraction of the signal. First, the 50Hz notch filter is used to remove the power frequency noise. McFarland et al. compared four kinds of spatial filtering technology to improve the SNR of EEG signal, and the conclusion shows the superiority of CAR and large Laplacian methods [21]; this paper uses CAR method as the spatial filter. The calculation of CAR is to subtract the average of all the electrodes from the selected channel. The formula for calculation is as follows:

$$V_i^{CAR} = V_i^{RAW} - \frac{1}{4} \sum_{j=1}^4 V_j^{RAW} \quad (1)$$

where V_i^{CAR} is filtered potential and V_i^{RAW} is the potential of the i electrode.

3.2. Empirical Mode Decomposition. EMD is used to stabilize the nonstationary signals and obtain IMF. The specific EMD decomposition process is given in document [22]. EEG signal can be decomposed into IMF components after the empirical mode decomposition, and the expression is as follows:

$$S(t) = \sum_{i=1}^N C_i(t) + R_n(t) \quad (2)$$

where $S(t)$ represents the original EEG signal, $C_i(t)$ is the i^{th} IMF, and $R_n(t)$ is residual components after screening.

Taking the 1-4s data of the F3 electrode channel in Figure 3, as an example, use EMD method to decompose

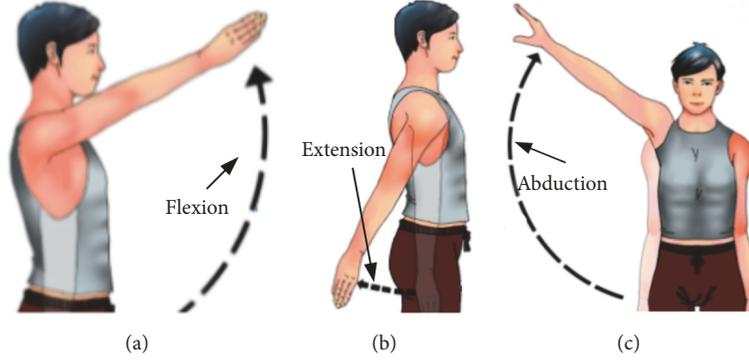


FIGURE 2: Three movements of shoulder joint: (a) flexion, (b) extension, and (c) abduction.

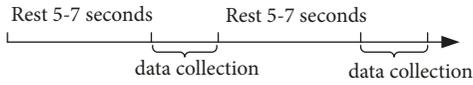


FIGURE 3: Timing for experimental process.

the denoised signal, and the IMF component is shown in Figure 4(a). Figure 4(b) is the AF domain information for the IMF component after FFT. From Figure 4(b), we can see that the μ and β rhythms of the motor imagery are mainly distributed in IMF1 and IMF2. In this paper, the sampled data of the selected 4 electrode channels are decomposed by EMD, respectively, do FFT to IMF1 and IMF2 of each electrode channel, and construct the i^{th} experiment EEG data matrix X_i ($i = N \times M$), where N corresponds to this article is 8 and M is the selected information of μ and β rhythms.

3.3. Common Spatial Pattern. The traditional CSP algorithm is essentially looking for a spatial filter to obtain more obvious eigenvectors after the signal passed through the filter, which makes a kind of signal variance reach a maximum, and another signal reaches a minimum. This method achieves the purpose of distinguishing two types of signals. Combining AF information of IMF with CSP to form AF-CSP and applying one versus one (OVO) strategy to AF-CSP make AF-CSP suitable for multiple classification problems, and the specific process is as follows:

$$R = \frac{X_i X_i^T}{\text{trace}(X_i X_i^T)} \quad (3)$$

where T is transpose operator. Then calculate the mixed space covariance matrix R_c of the two types motor imagery R_l , R_r as follows:

$$R_c = R_l + R_r \quad (4)$$

Eigenvalue decomposition for covariance of mixed space is as follows:

$$R_c = U_C A_C U_C^T \quad (5)$$

where U_C is the eigenvector matrix and A_C is the eigenvalue diagonal matrix. Whitening matrix P is calculated in

$$P = A_C^{-1/2} U_C^T \quad (6)$$

The whitening matrix causes eigenvalues of transformed matrix be equal one, so we calculate transformed covariance matrixes S_j and S_k in

$$S_j = P R_l P^T \quad (7)$$

$$S_k = P R_r P^T \quad (8)$$

After whitening, the matrixes S_j and S_k have the same eigenvector, and the following formula can be obtained after the eigenvalue decomposition:

$$S_j = B A_T B^T \quad (9)$$

$$S_k = B A_R B^T \quad (10)$$

The desired spatial filter is obtained by the upper form $W = (B^T P)^T$, and we can get a new data matrix after filtering by W : $Z_{N \times M} = W_{N \times M} X_i$. Feature vectors can be obtained by

$$f_p = \log \left(\frac{\text{var}(Z_p)}{\sum_i^{2m} \text{var}(Z_i)} \right) \quad (11)$$

where Z_p ($p = 1, \dots, 2m$). The dimension of f_p can not exceed N at most; in this paper, we set $m = 2$, so we can get a vector of four dimensions.

Combined with OVO strategy, eigenvectors constructed between every two types of actions are

$$f_1 = [F_1, E_1] \quad (12)$$

$$f_2 = [F_2, A_1] \quad (13)$$

$$f_3 = [E_2, A_2] \quad (14)$$

where f_1 represents the eigenvector obtained after AF-CSP transformation of flexion (F_1) and extension (E_1) of shoulder

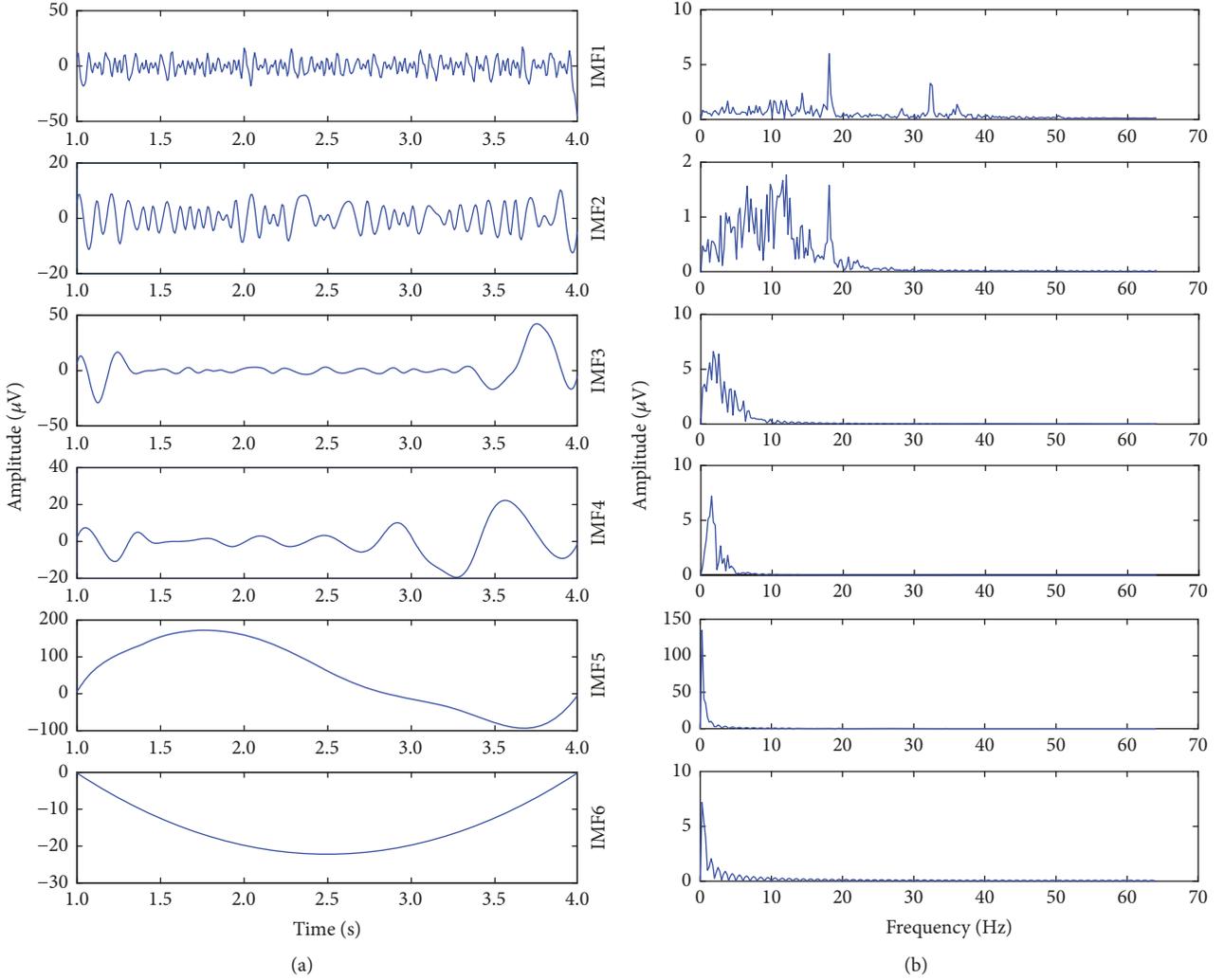


FIGURE 4: (a) IMF obtained from 3s EEG signals after EMD decomposition. (b) Amplitude-frequency domain information corresponds to per IMF of Figure 4(a).

joint. f_2 represents the eigenvector obtained after AF-CSP transformation of flexion (F_2) and abduction (A_1) of shoulder joint. f_3 represents the eigenvector obtained after AF-CSP transformation of extension (E_2) and abduction (A_2) of shoulder joint.

Figure 5 shows one example of the eigenvector f_1 constructed by AF-CSP of subject B, and f_1 is constructed by two MI tasks of shoulder joint flexion and extension. In Figure 5, the lateral axis represents the sequence number of experiments and the vertical axis represents the eigenvalue. We can see clearly from Figure 5 that the selected 4-dimensional feature vectors constructed by AF-CSP are distinctly distinguishable. The final eigenvector constructed by the AF-CSP method is

$$f = \begin{bmatrix} F_1 & F_2 \\ E_1 & E_2 \\ A_1 & A_2 \end{bmatrix} \quad (15)$$

3.4. Twin Support Vector Machine. TWSVM, which is developed on the basis of traditional SVM, is a new machine learning method [23]. For the two-classification problem, TWSVM constructs a hyperplane for every class of samples, so that each class sample is closest to its own hyperplane and far away from another hyperplane. TWSVM solves two-classification problem by solving a set of quadratic programming problems (QPPs), and SVM solves all classification problems by solving one QPP. This strategy makes TWSVM work 4 times faster than a standard SVM [24].

Combine OVO strategy with standard TWSVM to get OVO TWSVM, and the OVO TWSVM has a better classification performance than the OVO SVM [25]. For k class classification problem, the algorithm constructs a two-classification TWSVM subclassifier between any two classes of samples. Each subclassifier in OVO TWSVM needs only two classes of samples for training. Two hyperplanes are needed to train two types of sample i and j , such as

$$x^T w_{ij} + b_{ij} = 0 \quad (16)$$

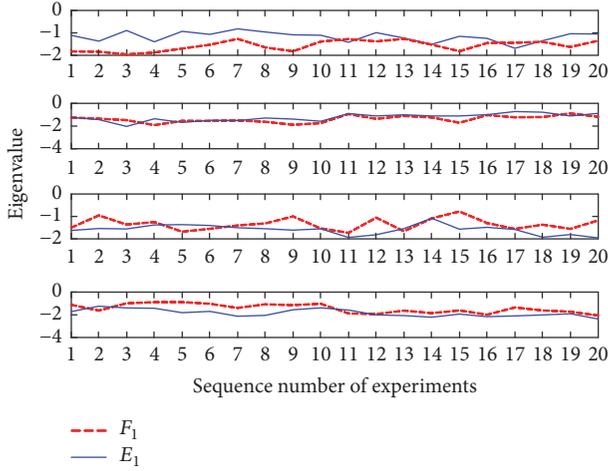


FIGURE 5: Twenty sets of four-dimensional eigenvectors f_1 obtained by AF-CSP method of flexion and extension of shoulder joint.

$$x^T w_{ji} + b_{ij} = 0 \quad (17)$$

where w_{ij} and w_{ji} are two normal vectors of hyperplanes and b_{ij} and b_{ji} are two hyperplanes. It is generally obtained by solving the following two-quadratic programming problem:

$$\min \frac{1}{2} \|A_i w_{ij} + e_{ij}^{(1)} b_{ij}\|^2 + \frac{c_{ij}}{2} e_{ij}^{(2)T} \xi_{ij} \quad (18)$$

$$\text{s.t. } (A_i w_{ij} + e_{ij}^{(2)} b_{ij}) + \xi_{ij} \geq e_{ij}^{(2)}, \quad \xi_{ij} \geq 0$$

$$\min \frac{1}{2} \|A_j w_{ji} + e_{ij}^{(2)} b_{ji}\|^2 + \frac{c_{ji}}{2} e_{ij}^{(2)T} \xi_{ji} \quad (19)$$

$$\text{s.t. } (A_j w_{ji} + e_{ij}^{(1)} b_{ji}) + \xi_{ji} \geq e_{ij}^{(2)}, \quad \xi_{ji} \geq 0$$

where c_{ij} is a penalty parameter and e_{ij} is a column vector of 1.

For the nonlinear separable phenomenon of training data, OVO TWSVM needs to solve the following optimization problems when training two samples of i and j :

$$\min \frac{1}{2} \|K(A_i, C^T) w_{ji} + e_{ij}^{(1)} b_{ji}\|^2 + \frac{c_{ij}}{2} e_{ij}^T \xi_{ij} \quad (20)$$

$$\text{s.t. } (K(A_i, C^T) w_{ji} + e_{ij}^{(2)} b_{ji}) + \xi_{ji} = e_{ij}^{(2)}, \quad \xi_{ij} \geq 0$$

$$\min \frac{1}{2} \|K(A_j, C^T) w_{ji} + e_{ij}^{(2)} b_{ji}\|^2 + \frac{c_{ji}}{2} e_{ij}^T \xi_{ji} \quad (21)$$

$$\text{s.t. } (K(A_j, C^T) w_{ji} + e_{ij}^{(1)} b_{ji}) + \xi_{ji} = e_{ij}^{(1)}, \quad \xi_{ji} \geq 0$$

Taking three-classification problem in two-dimensional space as an example to illustrate the process of OVO TWSVM with Figure 6. Taking i and j in (21) and (22) as 1 and 2 and solving them, we can get a two-class OVO TWSVM that classifies class 1 and class 2, i.e., subclassifier 1 in Figure 6. The subclassifier 2 and the subclassifier 3 can be obtained by a similar method, which constitute the OVO TWSVM for solving the three-classification problem together. Taking the green dot in the figure as an example to illustrate how to

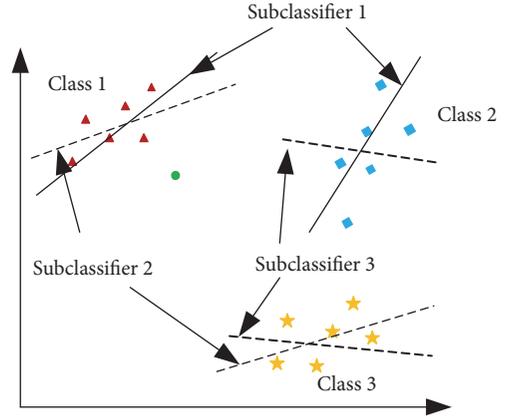


FIGURE 6: The sketch map of OVO TWSVM.

use OVO TWSVM. First, calculate the distance from the dot to the two hyperplanes of subclassifier 1; because the dot is close to the hyperplane of class 1, class 1 gets 1 vote. Second, calculate the distance from the dot to the subclassifier 2 in the same manner; since the dot is close to one hyperplane of subclassifier 2 in class 1, class 1 gets 1 more vote. Third, calculate the distance from the dot to the subclassifier 3; because the dot is close to one hyperplane of subclassifier 3 in class 2, class 2 got 1 vote and class 3 got 0 vote. OVO TWSVM classify the classified sample into class 1.

For the three-classification problem of this article, three OVO TWSVM subclassifiers need to set 6 penalty parameters: c_{11} , c_{12} , c_{21} , c_{22} , c_{31} , and c_{32} . This article sets $c_{11}=c_{21}=c_{31}$, $c_{12}=c_{22}=c_{32}$.

3.5. Multiobjective Genetic Algorithm. The core of the multi-objective genetic algorithm is to coordinate the relationship between the target functions and to make the target functions reach the Pareto optimal set. The quality of a solution in the Pareto optimal set is defined according to the dominance criterion. Any solution λ of the Pareto optimal set can be seen as an acceptable solution. If a solution λ_1 is no worse than λ_2 in all objectives and λ_1 is better than λ_2 in at least one objective, then we define λ_1 dominates λ_2 . The multiobjective optimization problem can be stated as

$$\text{Maximize } F(\omega) = \{f_1(\omega), f_2(\omega), \dots, f_n(\omega)\} \quad (22)$$

where $f_1(\omega), f_2(\omega), \dots, f_n(\omega)$ are the n objective functions and ω represents the parameters of the model.

The NSGA algorithm based on the fitness sharing technique proposed by Goldberg is based on the principle of non-dominated sorting to classify individuals in the population. And it can obtain a uniformly distributed Pareto optimal set or noninferior solution. However, the shortcomings of the algorithm are that the computational complexity is high and the sharing parameters need to be designated by human beings. Therefore, Deb's NSGA-II algorithm, which introduces fast nondominated sorting and elitist strategy to define crowding distance instead of fitness sharing, reduces the complexity of the algorithm and improves the computation

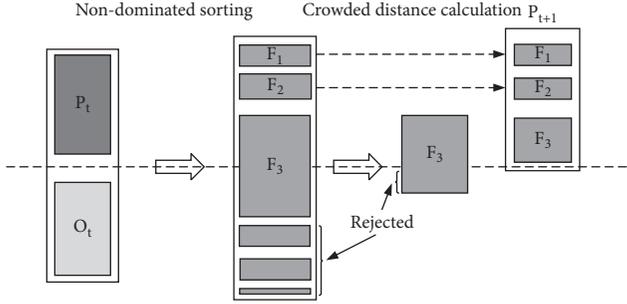


FIGURE 7: The steps of NSGA-II algorithm.

efficiency. NSGA-II overcomes three shortcomings of NSGA: the computational complexity which dropped from $O(MN^3)$ to $O(MN^2)$ (where M is the number of objectives and N is the population size), an elitist-preserving approach, and no sharing parameters which need to be specified. More details can be seen at [26].

Figure 7 describes the application of the dominance criterion in the NSGA-II algorithm. NSGA-II algorithm starts from an initialization population P_t and each individual in a population is no worse than the remaining individuals in the population. The following steps combine the NSGA-II with OVO TWSVM to optimize the classification results of OVO TWSVM. And then generate offspring O_t from P_t through binary tournament selection, crossover, and mutation. Once a foreign source is obtained, the algorithm will combine the current population and the current generation into a group and classify them according to the nondominated sorting and crowding distance. N optimal solutions can be obtained in the final set.

4. Analysis of Experimental Results

4.1. Construction of the Objective Function. Using NSGA-II to optimize the parameters, this paper uses the correct rate to construct the target function. The objective function is given as follows:

$$CRF = \frac{CNF}{TNF} \times 100 \quad (23)$$

$$CRE = \frac{CNE}{TNE} \times 100 \quad (24)$$

$$CRA = \frac{CNA}{TNA} \times 100 \quad (25)$$

$$CR = \frac{(CNF + CNE + CNA)}{(TNF + TNE + TNA)} \times 100 \quad (26)$$

CRF , CRE , and CRA represent the correct rate of flexion, extension and abduction, CNF , CNE , and CNA represent the correct action number of flexion, extension, and abduction, and TNF , TNE , and TNA represent the total number of flexion, extension, and abduction. CR represents the total correct rate.

4.2. Processing Steps. In the training phase, as the search process goes deep, the whole population tends to gather the

TABLE 1: NSGA-II preset parameters.

Parameter	Name	Preset value
P	Population size	100
CR	Crossover rate	0.9
MR	Mutation rate	0.1

global Pareto optimal set until the maximum evolutionary algebra is reached. The process of multiobjective optimization usually has the following steps:

(1) The dataset can be divided into a training set and a test set (50% for training, 50% for testing) or 5-fold cross validation, and in this article we use 5-folder cross validation in the whole analysis process.

(2) Change the parameters of the OVO TWSVM and run the target function.

(3) Set evolution algebra or stop criteria [27].

(4) Analyze the global optimal set to get the optimal parameters.

Table 1 shows the preset parameters of NSGA-II algorithm. The population size of the article is 100 and the crossover rate and mutation rate are 0.9 and 0.1.

Table 2 shows the range of penalty parameters and kernel function width for OVO TWSVM.

Table 3 shows the partial Pareto optimal set of the subject A and the corresponding OVO TWSVM model parameters using the NSGA-II algorithm.

Figure 8 shows the Pareto optimal fronts constructed by all nondominated solutions when seven subjects reach the optimal recognition rate. The mark points represent the optimum solution.

Figure 9 shows the evolutionary convergence curve of different subjects. It can be seen from the figure that five of the seven subjects achieved the highest recognition rate within 200 generations, they are subjects A, B, C, E, and F, and their accuracies are 91.66%, 95.00%, 90.00%, 85.00%, and 85.00%. All the subjects continued to increase the evolution algebra until the 600 generations. It was found that two subjects D and G had converged in the 200 generation, namely, 85% and 88.33%, respectively. The recognition rate of subjects F increased to 85% in the 400 generations and remained stable within 600 generations.

Figure 10 shows the accuracy of the OVO TWSVM classification using the linear kernel and the RBF kernel, respectively, along with the corresponding penalty parameters $c1$, $c2$ and kernel function width λ , and S represent different subject. As can be seen from Figure 10, OVO TWSVM based on RBF kernel achieved the highest recognition rate on six subjects, up to 95% on subject B. Subject D's highest recognition rate appeared in linear kernel based OVO TWSVM, but the recognition result was only 3.33% less than that of RBF kernel based OVO TWSVM. Therefore, this paper chooses RBF kernel based OVO TWSVM as the final recognition model.

4.3. Comparison with Other Methods. Literature [20] uses LS-SVM to classify the motor imagery EEG signals; this paper applies the model to our own dataset for classification. Fivefold cross validation for each person's data, the results,

TABLE 2: OVO TWSVM parameter range.

Parameter	Name	Lower limit	Upper limit
$c1$	Penalty parameter 1	2^{-3}	2^3
$c2$	Penalty parameter 2	2^{-3}	2^3
λ	RBF kernel function	2^{-20}	2^3

TABLE 3: Pareto optimal set of subject A and the corresponding model parameters.

$c1$	$c2$	λ	CRF	CRE	CRA	CR
0.2031	0.0100	0.7810	0.95	0.85	0.90	0.9000
5.0102	5.7356	0.1130	0.40	0.20	1.00	0.5333
5.4325	6.6230	0.2384	0.75	1.00	0.65	0.8000
6.7280	7.7008	0.7514	0.95	0.95	0.85	0.9166
4.1136	0.0100	0.6551	0.95	0.85	0.90	0.9000
4.8389	7.3435	0.2332	0.80	1.00	0.60	0.8000
3.6172	4.8072	0.1111	0.45	0.15	1.00	0.5333

TABLE 4: Comparison between AF-CSP and EMD-CSP.

Subjects	A	B	C	D	E	F	G	Mean
AF-CSP	91.66	95.00	90.00	85.00	85.00	85.00	86.67	88.57%±3.61%
EMD-CSP	88.33	88.33	81.66	76.67	88.33	83.34	73.34	82.85%±5.61%

and the corresponding parameter settings are shown in Figure 11. Where c is the penalty parameter, g is the kernel function width. As can be seen from Figure 11, LS-SVM with RBF kernel has the highest recognition rate of 92.09% on subject B, five other subjects were over 75%, and the worst recognition rate on subject F is 66.66%. Based on the above results, it can be considered that this method performs well on the data sets of different subjects.

We also use BP with momentum, ELM, PSO-SVM, and grid search OVO TWSVM to classify the datasets. The result of the classification is shown in Figure 12. It can be seen from Figure 12 that the proposed method (88.57%±3.61%) has the highest average recognition rate among the seven subjects compared to other recognition algorithms. LS-SVM (79.64%±7.47%), GS OVO TWSVM (79.99%±7.76%), and PSO-SVM (75.95%±7.86%) have similar average recognition effect on seven subjects, and the average classification results of BP (61.70%±6.62%) and ELM (61.42%±5.37%) in this dataset are lower than that of SVM classifier. This shows that SVM classifier has a great advantage in small sample machine learning.

Since the time domain signal is more similar, especially in the single joint motor imagery EEG signal, in this paper, we combine amplitude-frequency (AF) domain information with CSP, namely, AF-CSP to get feature vectors. Different from the traditional method of combining time domain signal with CSP to construct feature vector, AF information can reveal further the difference between different actions in frequency domain and amplitude range; thus, CSP can get stronger feature extraction ability. In this paper, we compare the classification rate using AF-CSP and CSP directly after EMD, as shown in Table 4. As can be seen from Table 3, the proposed method has achieved a higher recognition rate

among the 5 subjects, and it might be interpreted that AF-CSP is a more effective feature extraction method.

In order to distinguish the significance of the proposed method in this article, the one-way analysis of variance (ANOVA) method was used to compare with the other five algorithms. The p -value values are shown in Table 5, where $p1$, $p2$, $p3$, $p4$, and $p5$ represent the p -values between the proposed method and LS-SVM, BP, ELM, PSO-SVM, and grid search OVO TWSVM. When $p \leq 0.05$, there is a significant difference in the recognition effect between the two algorithms. It can be seen from Table 5 that the p -values between the proposed method and the five-other methods are significantly less than 0.01, which proves that the proposed algorithm has significant classification in improvement performance compared with other algorithms.

5. Discussion

Common spatial pattern is widely used in motor imagery to extract EEG features [28]. This method uses supervised learning to obtain two types of filter to separate two motor imagery tasks. In recent years, several methods like CSSP [29], RCSP [30], SSCSP [31], FERCSP [32], SBCSP [33], and FBCSP [34] have been used to improve conventional CSP. But the drawback of CSP is that it needs a lot of electrodes. AF-CSP, which taking into account the AF information in the EEG signals, uses only four electrode channels to achieve a better recognition effect than conventional CSP.

And an optimized OVO TWSVM using NSGA-II is used to improve the accuracy of motor imagery; the mean accuracy of the proposed method is 88.57%±3.61%. However, since there are no articles consistent with the contents of this paper, we can only discuss articles that are similar to our research

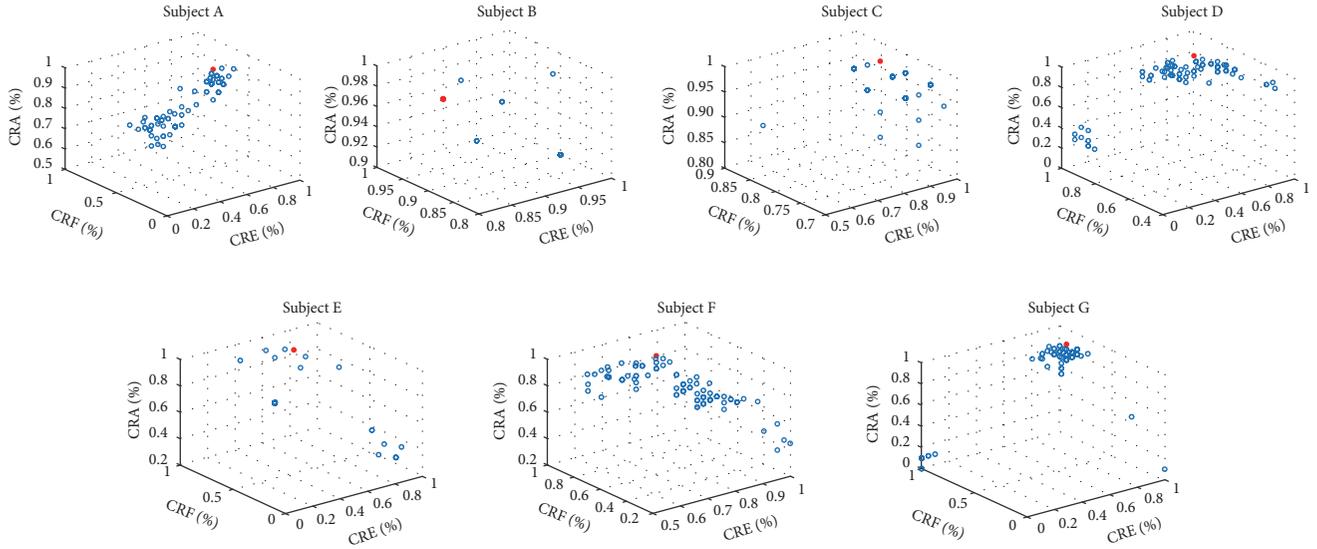


FIGURE 8: The Pareto optimal front of seven subjects based on RBF kernel OVO TWSVM.

TABLE 5: p -value score based on RBF kernel OVO TWSVM and five other methods.

Subject	$p1$	$p2$	$p3$	$p4$	$p5$
A	1.06e-05	3.90e-07	6.29e-08	2.78e-08	7.22e-05
B	3.80e-02	9.72e-06	6.97e-13	1.00e-03	5.33e-04
C	1.07e-07	6.27e-09	3.60e-11	3.55e-05	2.37e-07
D	2.84e-05	3.99e-07	9.81e-07	6.09e-07	3.00e-04
E	1.00e-04	3.64e-12	3.04e-06	1.84e-06	2.61e-11
F	3.88e-07	8.30e-08	9.84e-08	5.93e-06	3.40e-03
G	3.00e-04	3.82e-09	7.56e-06	4.91e-05	5.94e-07

TABLE 6: Mean confusion matrices of seven subjects for the proposed method.

	Flexion	Extension	Abduction
Flexion	90.71	7.57	3.50
Extension	5.50	82.14	5.79
Abduction	3.93	10.26	90.71

contents. Literature [18] points out that the manual extraction of features in the traditional biological signal pattern recognition model may produce information loss; thus, CNN with deep learning is introduced to identify of changes in the brain topographic map. The results show that three kinds of motor imagery in the hand are 65.51% and the kappa value in this experiment is 0.481. However, there is still a great deal of uncertainty about the intention recognition of single hand. In [19], the wavelet coefficients are calculated from EEG signals as feature and employed including quadratic discriminant analysis, naive Bayes quadratic, decision tree, K nearest neighbors, and SVM classifiers to identify shoulder and elbow joint movement. The highest recognition rate is 84.91%, using the SVM classifier. The weakness of this study is that it only recognizes the shoulder and elbow joint

movements of one hand and there is no further analysis of the complex movement of the single joint.

It should be noted that we use AF-CSP to extract features and enhance the feature extraction ability of CSP. OVO TWSVM is used to identify three types of upper limb movement; the lowest recognition rate is 85.00%. Compared with the current research, we have made a more in-depth analysis of single joint multiclass motor imagery. A more sophisticated analysis can be done using the mean confusion matrices given in Table 6 and the kappa value calculated by Table 6 is 0.82. It appears that the extension action is more difficult to distinguish with flexion and abduction of the shoulder joint using the proposed method. The flexion and extension of the shoulder joint obtain high recognition rate by the proposed method.

In addition, [35] proposes a novel correlation-based time window selection (CTWS) algorithm which considers the variation in the time latency during the MI task for MI-based BCI. CTWS adjusts the starting points of time window for both training and test samples using correlation analysis and shows significantly improvement than feature extraction algorithms without CTWS. As the feature extraction algorithms in the structure of the CTWS algorithm is substitutable, the focus of the next work is to combine CTWS

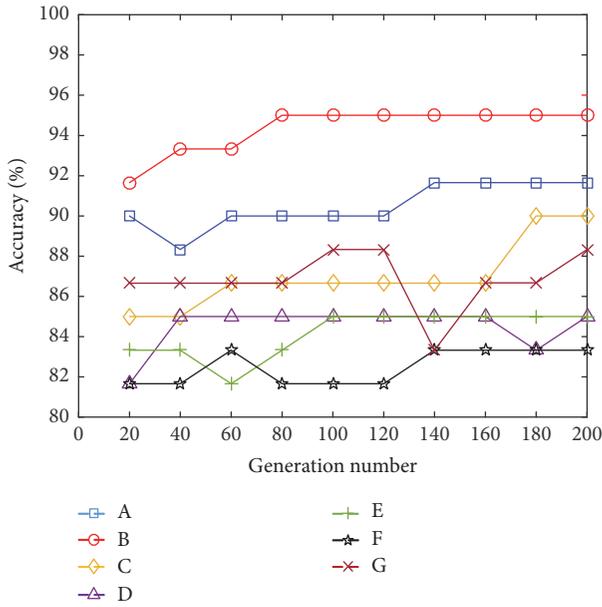


FIGURE 9: Evolutionary convergent fold line chart for different subjects.

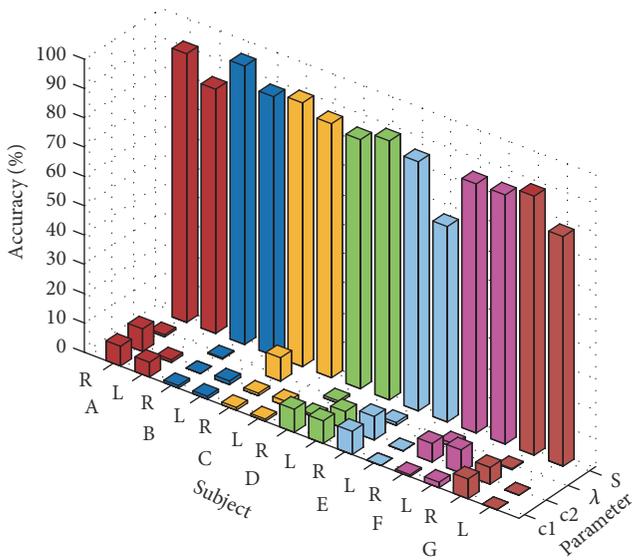


FIGURE 10: OVO TWSVM recognition effect diagram of different kernel function types.

with AF-CSP to increase the recognition rate of the extension and applying this model to the BCI systems.

6. Conclusion

This paper proposes AF-CSP replace the traditional EMD-CSP method. The main idea of this method is to analyze the μ and β rhythm information contained in each IMF component after EMD decomposition and extract corresponding AF information. This preprocessing method not only removes the influence of irrelevant frequency bands, but also strengthens the feature extraction ability of CSP. Secondly, this paper also uses NSGA-II to optimize the OVO TWSVM parameter

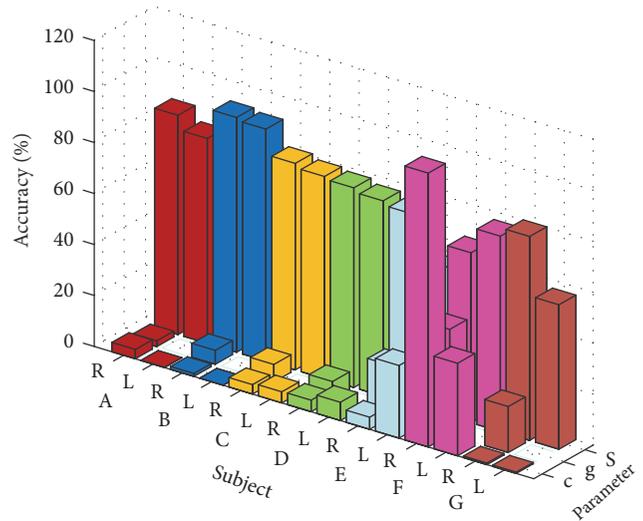


FIGURE 11: LS-SVM recognition rate diagram for different kernel functions.

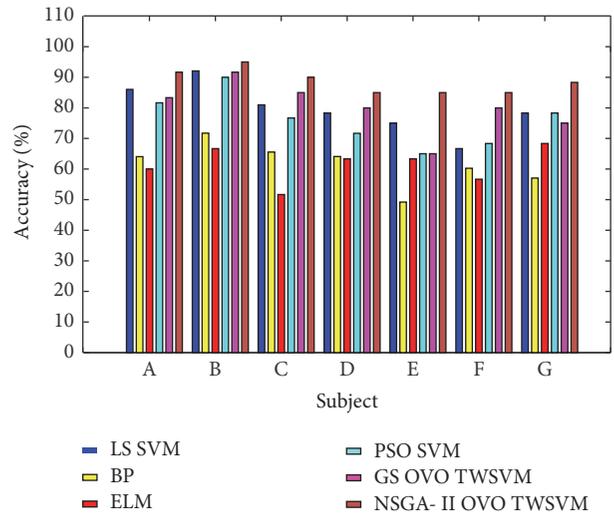


FIGURE 12: Comparison of recognition rates of different methods.

optimization process. Compared with other evolutionary strategies, the Pareto optimal set obtained by NSGA-II can make OVO TWSVM more robust. In the future, we may use this technology to evaluate the classification of real-time BCI and apply this method for reach and grasp tasks of a robotic arm. In order to improve the recognition rate of the method, we may use clustering or dimensionality reduction method to get a more obvious feature vector.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] N. Birbaumer, "Breaking the silence: brain-computer interfaces (BCI) for communication and motor control," *Psychophysiology*, vol. 43, no. 6, pp. 517–532, 2006.
- [3] J. Kalcher, D. Flotzinger, C. Neuper, S. Göllly, and G. Pfurtscheller, "Graz brain-computer interface II: towards communication between humans and computers based on online classification of three different EEG patterns," *Medical & Biological Engineering & Computing*, vol. 34, no. 5, pp. 382–388, 1996.
- [4] J. R. Wolpaw and D. J. McFarland, "Multichannel EEG-based brain-computer communication," *Electroencephalography and Clinical Neurophysiology*, vol. 90, no. 6, pp. 444–449, 1994.
- [5] B. Blankertz, C. Sannelli, S. Halder et al., "Neurophysiological predictor of SMR-based BCI performance," *NeuroImage*, vol. 51, no. 4, pp. 1303–1309, 2010.
- [6] N. Birbaumer, N. Ghanayim, T. Hinterberger et al., "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [7] J. D. Bayliss, "Use of the evoked potential P3 component for control in a virtual apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 113–116, 2003.
- [8] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [9] G. Pfurtscheller, "EEG Rhythms - Event-Related Desynchronization and Synchronization," in *Rhythms in Physiological Systems*, vol. 55, pp. 289–296, Springer, Berlin, Germany, 1991.
- [10] G. Pfurtscheller and F. H. L. da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [11] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.
- [12] D. Coyle, G. Prasad, and T. M. McGinnity, "A time-series prediction approach for feature extraction in a brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 4, pp. 461–467, 2005.
- [13] T. Kayikcioglu and O. Aydemir, "A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data," *Pattern Recognition Letters*, vol. 31, no. 11, pp. 1207–1215, 2010.
- [14] K. K. Ang and C. Guan, "EEG-Based Strategies to Detect Motor Imagery for Control and Rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 4, pp. 392–401, 2017.
- [15] P. Martín-Smith, J. Ortega, J. Asensio-Cubero, J. Q. Gan, and A. Ortiz, "A supervised filter method for multi-objective feature selection in EEG classification based on multi-resolution analysis for BCI," *Neurocomputing*, vol. 250, pp. 45–56, 2017.
- [16] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface," *Journal of Neuroscience Methods*, vol. 255, pp. 85–91, 2015.
- [17] C. Park, C. C. Cheong-Took, and D. P. Mandic, "Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 1–10, 2014.
- [18] W. Wang, S. Sun, C. Li, and Z. Tang, "Recognition of upper limb motion intention of EEG signal based on convolutional neural network," *Journal of Zhejiang University*, vol. 7, no. 51, pp. 1381–1389, 2017.
- [19] R. Rinku, A. Konar, and N. Tibarewala, "Classification of left-right arm and shoulder joint- elbow joint motor imagery data," in *Proceedings of the National Conference on Telecommunication & Control Engineering*, 2012.
- [20] S. Taran, V. Bajaj, D. Sharma, S. Siuly, and A. Sengur, "Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications," *Measurement*, vol. 116, pp. 68–76, 2018.
- [21] X. Tang, N. Zhang, J. Zhou, and Q. Liu, "Hidden-layer visible deep stacking network optimized by PSO for motor imagery EEG recognition," *Neurocomputing*, vol. 234, pp. 1–10, 2017.
- [22] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A: mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [23] R. Khemchandani and S. Chandra, "Twin support vector machines for pattern classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 5, pp. 905–910, 2007.
- [24] S. Ding, J. Yu, B. Qi, and H. Huang, "An overview on twin support vector machines," *Artificial Intelligence Review*, vol. 42, no. 2, pp. 245–252, 2014.
- [25] D. Tomar and S. Agarwal, "A comparison on multi-class classification methods based on least squares twin support vector machine," *Knowledge-Based Systems*, vol. 81, pp. 131–147, 2015.
- [26] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [27] A. R. J. Clark and R. M. Everson, "Multi-objective learning of Relevance Vector Machine classifiers with multi-resolution kernels," *Pattern Recognition*, vol. 45, no. 9, pp. 3535–3543, 2012.
- [28] Z. J. Koles, M. S. Lazar, and S. Z. Zhou, "Spatial patterns underlying population differences in the background EEG," *Brain Topography*, vol. 2, no. 4, pp. 275–284, 1990.
- [29] S. Lemm, B. Blankertz, G. Curio, and K.-R. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1541–1548, 2005.
- [30] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355–362, 2011.
- [31] Y. Shin, S. Lee, J. Lee, and H.-N. Lee, "Sparse representation-based classification scheme for motor imagery-based brain-computer interface systems," *Journal of Neural Engineering*, vol. 9, no. 5, Article ID 056002, 2012.
- [32] S. Yuxi, L. Yali, and W. Shengjin, "Filter ensemble regularized common spatial pattern for EEG classification," in *Proceedings of the Seventh International Conference on Digital Image Processing (ICDIP '15)*, vol. 9631, International Society for Optics and Photonics, 2015.

- [33] Q. Novi, C. Guan, T. H. Dat, and P. Xue, "Sub-band common spatial pattern (SBCSP) for brain-computer interface," in *Proceedings of the 3rd International IEEE/EMBS Conference on Neural Engineering (CNE '07)*, pp. 204–207, IEEE, Kohala Coast, Hawaii, USA, May 2007.
- [34] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers in Neuroscience*, vol. 6, 9 pages, 2012.
- [35] J. Feng, E. Yin, J. Jin et al., "Towards correlation-based time window selection method for motor imagery BCIs," *Neural Networks*, vol. 102, pp. 87–95, 2018.



Hindawi

Submit your manuscripts at
www.hindawi.com

