



Parallel genetic algorithm based common spatial patterns selection on time–frequency decomposed EEG signals for motor imagery brain-computer interface

Tian-jian Luo ^{*}

*College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China
Digital Fujian Internet-of-thing Laboratory of Environmental Monitoring, Fujian Normal University, Fuzhou 361005, China*



ARTICLE INFO

Keywords:

Motor imagery
Brain computer interface
Common spatial pattern
Time segments and frequency bands
Parallel genetic algorithm

ABSTRACT

Since the nonlinear and non-stationary characteristics of electroencephalogram (EEG) signals, motor imagery based brain-computer interface (MI-BCI) have problems of poor recognition accuracy and robustness across subjects and recording sessions. To address such problems, we extract common spatial pattern features on a detailed decomposition of time segments and frequency bands (TSFBCSP) on EEG signals, and propose a novel encoding approach of genetic algorithm (GA) for features selection. The proposed approach selects the most robust and discriminative CSP features to enhance the accuracy and robustness of EEG signals recognition. To improve the efficiency of TSFBCSP features selection, the parallel GA is implemented on a MapReduce framework (MRPGA). Comparative experiments have been done on two publicly available datasets (2a and 2b) from BCI competition IV. The excellent average recognition accuracy shows that the proposed TSFBCSP-GA/MRPGA has a promising candidate performance improvement on MI-BCI. Classification accuracies based on the optimal EEG segments and sub-bands have achieved 74.92 % and 81.04 % for the dataset 2a and 2b, respectively, and the faster classification accuracies of 74.54 % and 79.78 % has been achieved for the MRPGA architecture. Ablation study further proves the feasibility and effectiveness of the proposed algorithm, and discusses the parameters setting of applying it to actual MI-BCI applications.

1. Introduction

Brain computer interface (BCI) [1–3] is a recent promising technology that establishes a signal transmission path between brain and computer, and users can directly control external devices through brain and bypass the conduction path between brain and muscle system. Therefore, BCI technology has broad application prospects in patients with paralysis and atresia, as well as in the aerospace and extreme environment [4–6]. BCI technology involves the intersection of signal processing, machine learning, cognitive neuroscience, and other disciplines [7–9]. The theory of BCI technology is based on electrophysiological and hemodynamic of human brain. Based on such two theories, the activations of brain using BCI can be measured by the electroencephalography (EEG) [10], electrocorticography (ECoG) [11], and magneto-encephalography (MEG) [12]. Among them, the low cost and high time resolution of EEG signal make it the preferred signal means for BCI technology.

The non-invasive multivariate EEG signals are out of requirements of ethical constraints and invasive brain surgery, and have become a more suitable way to construct BCI for normal human beings [13]. For the multivariate EEG based BCI, sensorimotor rhythm [14], event-related potential [15], and steady-state visual evoked potential [16] are three main paradigms for applications. Motor imagery (MI) is a common way to evoke sensorimotor rhythm using an autonomous way by human beings [17]. The motor imagery tasks suppress the power of mu and beta rhythms on the contra-lateral sensorimotor regions of brain, and such inhibitions called event-related desynchronization (ERD) [18]. Based on the electrophysiological and hemodynamic theory, the suppressions of ERD are widely distributed among all brain regions. In addition, due to the topological structure of motor neurons, the EEG signals collected by the cerebral cortex are usually mixed by multiple sensorimotor neurons [19], resulting in a poor spatial resolution of the raw EEG signals and degrading the pattern recognition performance.

To improve the spatial resolution of ERD suppressions and ensure the

* Address: College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350117, China.
E-mail address: createopenbci@fjnu.edu.cn.

performance of pattern recognition, the common feature extraction methods include spectral analysis, auto-regression, source reconstruction and common spatial pattern (CSP) [20]. Among them, the CSP features have been widely applied to MI-BCI. By solving the optimal spatial filters, such algorithm was used to maximize the variance of one MI task and minimize the variance of another MI task. Therefore, the CSP algorithm was suitable for multivariate EEG signals feature extraction. In fact, the optimal spatial filters extracted ERD suppressions embedded in EEG signals by selecting or re-weighting signals recorded on multiple electrodes. To regularize the over-fitting on conventional CSP features, the most popular way to enhance CSP features was to extract spatial features while considering frequency bands. The most representative researches included common spatio-spectral patterns (CSSP) [21], sub-band/filter bank CSP (FBSCP) [22], regularized FBSCP (FBRCSP) [23], discriminative FBCSP (DFBCSP) [24], sparse FBSCP (SFBCS) [25], and local region frequency CSP (LRFSCP) [26]. These improved CSP features conducted different regularization strategies while extracting spatial features on frequency bands. For example, the LRFSCP considered the local region frequency optimized spatial features from individual-specific channels and filter banks.

Different from conducting regularizations on frequency bands, other researches considered both sequential relationships and frequency bands to enhance CSP. The most representative researches included separable CSSP (SCSSP) [27], spectrally weighted CSP (SWCSP) [28], penalized time-frequency band CSP (PTFBCSP) [29], sliding window with longest consecutive repetition based CSP (LCR-SW-CSP) [30], temporally constrained sparse group spatial pattern (TSGSP) [31], and time-frequency CSP (TFCSP) [32]. These enhanced CSP series features conducted penalizations on both filter banks and time segments or windows while extracting spatial features. For example, the TSGSP algorithm first decomposed EEG signals into multiple spectrum-specific components, and the filter banks were put on a set of sliding time windows of such components. Then, a joint sparse optimization on both filter banks and sliding windows was used to extract regularized and robust CSP features.

Nevertheless, the existed regularization and enhancement of CSP on frequency bands and sequential relationships needed a lot of computational complexity and complex mathematical transformation. The best acceptable regularizations were commonly used Tikhonov regularization, probabilistic regularization [33], or projected EEG covariance matrix into Riemannian space and performed classification using a minimum to the mean of Riemannian distance [34]. Instead of using mathematical regularization for selecting features, there were a lot of feature selection methods for motor imagery BCI using computational intelligence algorithms. During features extractions phase on MI-BCI, the conventional computational intelligence algorithms were commonly used for selecting appropriate combinations of electrodes, so as to improve the accuracy of pattern recognition of MI-BCI.

To select a subset of relevant and non-redundant electrodes for effective motor imagery classification, researchers adopted a combination of graph theoretic spectral and quantum genetic algorithm (GA) [35]. The other researchers used a modified particle swarm optimization (PSO) algorithm called neuro-evolutionary process to select the optimal electrodes, and the conventional computational intelligence algorithms were also compared for the improvements of motor imagery classification [36]. Another research used a novel NSGA-II algorithm extended from GA with a hierarchical individual representation to exclude the irrelevant channels for motor imagery [37]. The ensemble extreme learning machines (EELMs) were also applied to classify motor imagery EEG signals, and the PSO was used to simultaneously optimize weights and biases of EELMs [38]. Combining Hilbert-Huang transform and the CSP features, the GA was used for selected a subset of the extracted composite features to improve the classification accuracy [39]. Another research used a hierarchical GA for selection of MI relevant features to estimate the time-resolved strength of motor imagery representation [40].

To sum up, the existing computational intelligence algorithms were used for feature extraction of ERD suppression, which were generally used for optimal electrodes combination selection, optimal classifier parameter acquisition, and optimal classical CSP features weighting. However, there were no relevant researches on the selection of appropriate time segments and frequency bands in the process of solving CSP features, and most of the current feature selection methods by computational intelligence algorithms were lack of considering the efficiency of classification for the MI-BCI.

The classical MI paradigm presented a cue to guide the imagine direction of motors for subjects. The commonly used paradigm presented a cue for motor imagery, and asked the subject to image the direction of motors followed the cue. Although we could determine the exact time of the cue's exhibition during EEG signal recording, due to different reaction time, we cannot accurately judge the occurred time of motor imagery for an individual-specific subject [41]. For the classical MI paradigm, the range of [0, 1] second was regarded as the preparation phase, and the following rang of [1,4] seconds was treated as the performing phase. Therefore, to obtain more discriminative ERD suppressions, the appropriate time window covering the whole MI process should be selected, and the MI preparing or MI ending time segments should be removed as much as possible. When the cue for MI is presented, individuals have different time durations for preparing MI and performing MI. However, based on the heuristically instruction of BCI competition, the majority algorithms selected the best time segment and frequency band as [0.5, 2.5] seconds and 7–30 Hz during 4 s motor imagery procedure [42]. However, due to the nonlinear and non-stationary characteristics of EEG signals, the existing enhancement schemes of CSP algorithm cannot cope with cross-sessions and cross-subjects in MI.

The different segments and rhythms of EEG signals reflect the current situation of the subjects. In common, the passive or active stimuli have been applied to the subjects, such as auditory, visual, motor imagery, and affective stimulus, and the different lengths of EEG segments under different frequency bands of EEG rhythms are recorded from the brain to reflect the situations of such stimuli [43]. For the affective BCI, EEG signals were recorded during applying emotion-inducing video clips to subjects, and such EEG signals were adopted for the emotion classification. Some researchers have analyzed different EEG rhythms. By classifying asymmetry measurements on different rhythms of delta (0–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz), and gamma (32–64 Hz), an effect excellent classification result has been obtained [44]. Meanwhile, the researchers also analyzed different EEG segments. By statistical analysis of (0–600 ms) and (0–1200 ms) EEG segments, they found that a longer EEG segment will obtain better classification results [45]. For the studies of auditory BCI induced by musical stimuli, researchers used a frequency band of (42–49 Hz) to extract features from polyphonic music, and a segment of 1400 ms EEG segments are used for classification [46]. More researches on the aesthetic preference of visual stimuli have also used different EEG segments and rhythms. For a research of Chinese dance posture esthetic preference, researchers extract EEG signals of (0–4000 ms) on a frequency band of (0.15–30 Hz) for the preference dancer classification [47]. Another research of Chinese typeface implicit esthetic preference, researchers recorded the EEG signals of (150–200 ms), (200–300 ms), (300–400 ms), and (400–600 ms) to under a frequency band of (0.15–30 Hz) to measure the esthetic preference [48]. The effective, auditory, and visual paradigms of BCI are the passive BCI, without loss of generality, the active BCI of motor imagery also should consider the EEG segments and frequency bands of EEG rhythms during classification.

At present, there are some researches considering the EEG segments and frequency bands of EEG rhythms during motor imagery classification. In Refs. [31,32], researchers decomposed MI process into reaction, action, and completion stages, filtered each stage and extracted CSP feature, and combined and regularized all CSP features for pattern recognition. However, the decompositions for time segments and

frequency bands of the existing methods were too rough. Since the tremendous variations of non-stationary EEG signals on different time-domain and frequency-domain, a rough decomposition will lose a lot of spatial features. To obtain the most comprehensive and detailed spatial features, this paper decomposed the time segments and frequency bands of EEG signals at the minimum units (0.1 s with 1 Hz), and extracted CSP features on the decomposed time segments and frequency bands, called TSFBCSP features. The comprehensive and detailed decompositions will bring a large number of MI non-related redundant features, which will degrade the performance of pattern recognition.

To alleviate the influence of redundant features from TSFBCSP features, we introduced the GA [49] to select optimal features for specific subject, and confirmed that each session selected sufficient and appropriate spatial features. Meanwhile, to improve the efficiency of feature selection by GA, we took the classification accuracy as the optimization object to construct fitness function, and used the MapReduce framework [50] to build parallel genetic algorithm (MRPGA). After multiple rounds of iterations, the introduced GA/MRPGA algorithm can not only select subject-specific optimal features to improve the accuracy of MI-BCI classification, but also reduce the feature dimensions for different subjects and suppress the negative impact of noises.

The contributions of this paper are listed as follow:

(1) This paper extracted TSFBCSP features from EEG signals by the overlapped time segments and frequency bands. Based on the decompositions of 0.1 s and 1 Hz respectively, the rapidly varying ERD suppressions of MI hidden in nonlinear and non-stationary EEG signals can be extracted for classification.

(2) To improve the classification efficiency and alleviate the MI non-related redundant features, a MapReduce framework based parallel GA was introduced to select MI related features for a faster, precise, and robust classification of MI tasks.

(3) To validate the feasibility and effectiveness of the proposed TSFBCSP-MRPGA algorithm, two public benchmark MI-EEG datasets of nine subjects were selected to perform classification experiments, and the proposed algorithm showed precise and robust results based on the trade-off of performance and efficiency.

The rest of the paper is organized as follow: Section 2 describes the TSFBCSP features extracted in the paper. Section 3 introduces the chromosome coding method, fitness function definition, GA for TSFBCSP features selection and the extension of MapReduce based parallel GA. Section presents the experimental results and ablation studies. Section 5 concludes the paper and points out the future works.

2. Time segment with frequency bands common spatial patterns

2.1. Time segments decomposition for multivariate EEG signals

In majority time segmentation of 4 s, motor imagery tasks select [0,4] seconds, and such selection was validated by a lot of researchers [20,22,28]. However, the characteristics of multivariate EEG signals on a large sampling rate reflect ERD variations in a teeny time range. To solve this uncertainty on time domain, some researchers explored the time stages on EEG signals, and analyzed the ERD variations on different time stages. For example, the TFCSP algorithm decomposed the time domain to three stages: preparation stage ([0, 1] second), imagination stage ([1,3] seconds), and post-imagination stage ([2,4] seconds); the TSGSP algorithm decomposed the time domain to a set of time segments decomposed by 0.5 s: [0, 2] seconds, [0.5, 2.5] seconds, [1,3] seconds, [1.5, 3.5] seconds, and [2,4] seconds. However, the range of the existing “stage decomposition” or “time segment decomposition” is too rough. Since the ERD suppressions varied quickly in EEG signals during motor imagery, a rough decomposition will lose some key features for pattern recognition. Due to the above two time domain decomposition algorithms do not have feature selection process, a more detailed decomposition of time domain will matter a large time complexity.

This paper introduced a feature selection algorithm before recog-

nizing motor imagery, so we can perform a comprehensive decomposition of time domain to represent detail variations of ERD suppressions. In addition, the feature selection algorithm also prevented redundant information that non-related to motor imagery, such as artifacts and noises. In particular, we assumed that the raw multivariate EEG signals as $x \in R^{N \times T}$, where N and T denote the number of recording electrodes and the number of recording sampling points. The majority researches used a range of 2 s as each time segment, so we follow this setting to decompose the time segments as follow:

$$x_I = x(:, d \times s : (d + T/2) \times s), d = 0.1, 0.2, \dots, 2 \quad (1)$$

where $x_I \in R^{21 \times N \times t}$ represent the time segments set after decomposition from raw multivariate EEG signals. In the experiment, we decomposed the original EEG signals to 21 time segments, and the length of each time segment was $t = \frac{T}{2}$.

Motor imagery belongs to active brain computer interface. Researchers have shown that when to start or end motor imagery tasks and how long the motor imagery tasks carried out have individual differences. It is a difficult challenge to obtain the actual time domain process of different stages for different motor imagery subjects. Therefore, the time segments decomposition of this paper adopted a longer overlap window to take all situations and stages into account for the process of motor imagery, including imagination preparation stage, imagination stage, and post-imagination stage. However, this approach will bring non-related information for pattern recognition. These redundant features will be eliminated in the subsequent features selection process.

2.2. Frequency bands decomposition for multivariate EEG signals

Multivariate EEG signals contained three main features: time domain, frequency domain, and spatial domain. Different from time domain and spatial domain, since the process of motor imagery caused power variations in specific frequency bands, including μ rhythm (8–12 Hz) and β rhythm (18–24 Hz), it is especially important to perform frequency bands decomposition before feature extraction. To obtain stable ERD suppressions in a specific frequency band, the classical filter usually used a range of [4,40] Hz. However, the non-stationary characteristics of multivariate EEG signals resulted in different frequency bands across subjects and sessions for different motor imagery tasks. To solve the unstable frequency bands, the classical FBCSP algorithm introduced the filter bank configuration, and decomposed the entity frequency band of [4,40] Hz into a set of sub-bands: [4,8] Hz, [8,12] Hz, ..., [36,40] Hz. However, such range of frequency bands decomposition was also too rough. A rough frequency bands decomposition will still cause some key ERD suppressions to be lost, since the multivariate EEG signals were recorded with a high sampling rate.

Based on the classical filter bank, this paper introduced an overlapping frequency bands decomposition strategy and adopted a more detail frequency bands decomposition. To confirm which spacing of the frequency bands division was the best one, a preliminary experiment was conducted on the parts of EEG signals from the selected datasets. The results have shown that a division interval of 1 Hz and 0.5 Hz achieved almost the same classification accuracy, which was higher than the conventional division interval of 2 Hz or 3 Hz. However, the division of 0.5 Hz obtained more components in TSFBCSP features, which increased the time complexity in feature selection and classification, but only achieved a bit little improvement of classification performance. For a comprehensive consideration, a 1 Hz spacing of frequency decomposition among [4,40] Hz is selected in this paper. Fig. 1 illustrates the diagram of the time segments and frequency bands decomposition from multivariate EEG time-series.

In consequence, we also followed a range of 4 Hz in each frequency band, and adopted an 1 Hz spacing for the overlapping frequency bands for band-pass filtering: [4,8] Hz, [5,9] Hz, ..., [36,40] Hz. This frequency bands decomposition strategy can satisfy more frequency band

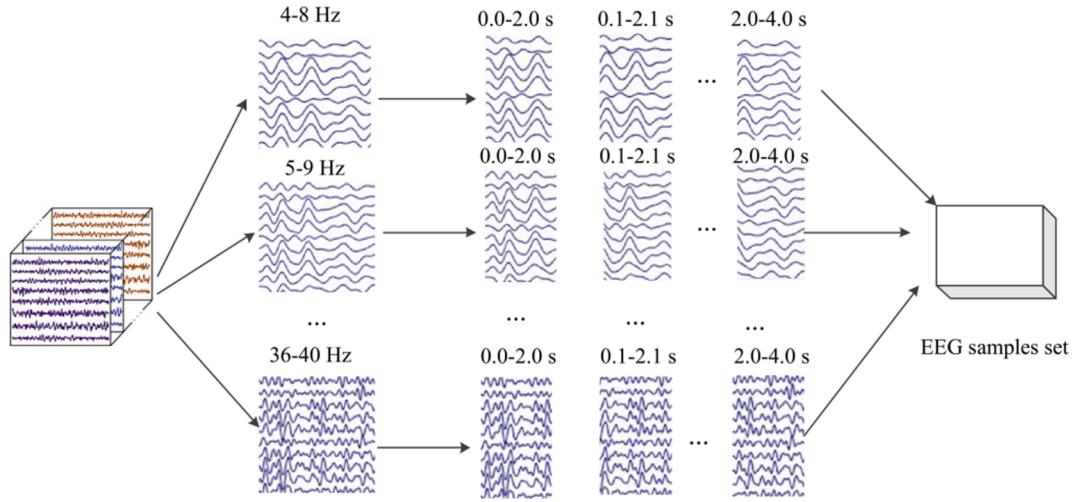


Fig. 1. The diagram of the time segments and frequency bands decomposition from multivariate EEG time-series.

changes in cross-subject and cross-session, and all variations in ERD suppressions were taken into account in the subsequent feature extraction. The band-pass filters for extracting all overlapped sub-band components ($x_i^j \in R^{33 \times N \times t}$) from each multivariate EEG signal time segments x_i were zero-phase Chebyshev Type I Infinite Impulse Response (IIR) filters. The filtering was implemented using the `filtfilt()` function in MATLAB R2018b (MathWorks, Inc.). Similarly, the redundant non-related features will be ignored by the following features selection. The time segments and frequency bands decomposition algorithm is summarized in **Algorithm 1**.

Input:	Algorithm 1: The time segments and frequency bands decomposition algorithm Multivariate EEG time-series
Output:	The EEG samples set under time segments and frequency bands decomposition
Steps:	(1) Setting parameter of time segments decomposition; (2) Setting filter bank of [4,8] Hz, [5,9] Hz, ..., [36,40] Hz of frequency bands decomposition; (3) Apply the bank of Chebyshev Type I IIR filters to band-pass filtered the input multivariate EEG time-series; (4) Extract motor imagery trial (4000 ms) from each filtered multivariate EEG time-series by the onset of each cue; (5) Segmentation each motor imagery trial to time segments by Eq. (1) to form the EEG samples set, and each time segment share the label of the trial; (9) return EEG samples set.

In summary, after decompositions of time segments and frequency bands, the features of a single multivariate EEG signals after segmented and filtered were obtained, which contained a total of 693 (21×33) samples. In the follow-up, the individual-specific and session-specific MI-related features will be selected for pattern recognition.

2.3. Common spatial patterns

The ERD suppressions reflected by MI on the opposite brain regions cannot be obtained directly from multivariate EEG signals. The most popular feature extraction algorithm was common spatial pattern (CSP) [22,42]. Taking the binary MI tasks recognition as an example, the classical CSP algorithm constructed several spatial filters through optimization. The optimization process ensured that the variance of EEG signals for one MI task was maximized, and the variance of EEG signals for the remaining MI task was minimized. CSP algorithm had different implementation versions. This paper adopted the CSP algorithm to construct the optimal spatial filters on the constructed EEG time segments filtered on frequency bands. For the segmented and filtered EEG

signals $x_{i,p}^j \in R^{N \times T}$, where $x_{i,p}^j$ represented the i -th EEG time segment and filtered on the j -th frequency band, and the ground-truth category (MI task) belonged to p . First, the CSP algorithm computed the covariance matrix of $x_{i,p}^j \in R^{N \times T}$:

$$c_{i,p}^j = \frac{x_{i,p}^j x_{i,p}^{j,T}}{\text{tr}(x_{i,p}^j x_{i,p}^{j,T})} \quad (2)$$

Assumed that the categories of two MI tasks were $p = 1, 2$. The CSP algorithm solved the following object function to obtain optimal spatial filters:

$$\max J(w_i^j) = \frac{w_i^{j,T} c_{i,1}^j w_i^j}{w_i^{j,T} (c_{i,1}^j + c_{i,2}^j) w_i^j} \quad (3)$$

where spatial filters $w_i^j \in R^{N \times 1}$ will maximize $c_{i,1}^j$ and minimize at the same time $c_{i,2}^j$. The Rayleigh quotient form of equation (3) can be transformed into the following general eigen-value form:

$$c_{i,1}^j w_i^j = \lambda_i^j (c_{i,1}^j + c_{i,2}^j) w_i^j \quad (4)$$

where λ_i^j and w_i^j represented the generalized eigen-value and eigen-vector. To maximize and minimize, the obtained eigen-values were sorted descending. Then, the eigen-vectors corresponding to the first m eigen-values were taken as the spatial filter for category $p = 1$, and got the eigen-vectors corresponding to the last m eigen-values as the spatial filter for category $p = 2$ at the same time. In this way, the final CSP spatial filter was represented by $W_i^j = [w_{i,1}^j, w_{i,2}^j] \in R^{N \times 2m}$. The common used CSP features generally used the normalized covariance form as the final features f_i^j :

$$f_i^j = \log(\text{diag}(W_i^{j,T} c_i^j W_i^j / \text{tr}(c_i^j))) \quad (5)$$

where $\text{diag}()$ represents the diagonal operation, and $\log()$ represents the logarithmic operation. However, the above computation ignored the sequential relationships of multivariate EEG signals in the time domain. Thus, this paper used the following way to obtain CSP features on time-series:

$$f_i^j = \log((W_i^{j,T} x_i^j)^2) \quad (6)$$

In this computation, we acquired the power of multivariate EEG signals by a quadratic operation, and the logarithm operation was used to ensure the power variations of ERD suppressions within a certain range [18]. The classical CSP algorithm carried out on two categories.

For the multi-categories MI tasks, we embraced the one vs rest strategy to extend the CSP algorithm. We selected one category $asp = 1$, and the remaining categories were treated $asp = 2$. Hence, a k -categories MI tasks will take k rounds of CSP features extraction to constitute the TSFBCSP features. The TSBFCSP features extraction algorithm is illustrated in **Algorithm 2**.

Input:	Algorithm 2: The TSFBCSP features extraction algorithm EEG samples set
Output:	TSFBCSP features
Steps:	(1) Setting parameter of selected eigen-vector columnsm; (2) Compute average covariance matrix for the EEG samples of each class by Eq. (2); (3) Solve the general eigen-value decomposition of Eq. (4); (4) Compute the time-series form of TSFBCSP features by Eq. (6); (5) return TSFBCSP features

3. TSFBCSP features selection via GA/PGA

3.1. Genetic algorithm for TSFBSCP features selection

To exclude non-related features, the common used feature selection algorithms included: mutual information (MUI) [42], Kullback-Leibler divergence (KLD) [51], concave minimization (FSV) [52], and five-stage decoding of EEG (FSDE) [53]. However, these existing feature selections basically utilized unsupervised strategies or classifier regularizations, which have a bottleneck of recognition performance. This method employed a supervised TSFBCSP feature selection method by using the genetic algorithm (TSFBCSP-GA) [49]. The feature selection methods first genetically encoded a large number of redundant features, and then constructed the fitness function for GA using the cross-validation classification accuracy. To improve the efficiency of TSFBCSP-GA, a parallel GA was designed on MapReduce framework [50] for a faster feature selection (TSFBCSP-MRGA).

After feature extraction, a total 693 samples of TSFBCSP features were genetically encoded as integer coding chromosomes. By using such chromosomes on GA, the optimal TSFBCSP features for time segments and frequency bands were selected for the individual-specific and session-specific MI tasks with parameter-free process. GA is an intelligent search algorithm inspired by the biological world. It does not depend on the problem itself. All it needs is to evaluate the chromosomes generated by the algorithm through a pre-set fitness function to make the chromosomes with high fitness get a better chance of reproduction. Two termination conditions were given to GA: reaching the maximum iterations number or reaching the best fitness function.

Specifically, we encoded the 693 samples to the genes of number 1–693. During the iterations of GA, the genes were randomly selected by the parameters to constitute each chromosome. One generation of GA contained three steps:

(1) **Selection:** The roulette selection method was adopted, that is, the probability of selected chromosomes was proportional to the values of fitness function. Assumed the chromosome population size encoded by the TSFBCSP feature was g , and the fitness function of single chromosome x_i was $f(x_i)$. Then, the probability of selected chromosome x_i can be computed as:

$$P(x_i) = \frac{f(x_i)}{\sum_{j=1}^g f(x_j)} \quad (7)$$

The numerator was the fitness value of the individual chromosome x_i , and the denominator was the sum of fitness values for all chromosomes. Using the roulette selection method can ensure that the chromosomes with higher fitness values have a greater probability of being retained in the next generation, while chromosomes with smaller fitness values were easily eliminated. The chromosomes fitted for the specific subject in a specific session can be selected.

(2) **Crossover:** According to the crossover probability P_c , two chromosomes that were paired with each other will exchange parts of their

genes in a certain way to form two new chromosomes. During the crossover process, a random pair-wise paring strategy was used for crossover operation, and the restricted condition was to ensure that the number of offspring in range of $2 \leq g \leq k$, where k was a constant threshold. In addition, a new judgment criterion was added to prevent newly generated chromosomes from being the same as the parent, so as to avoid the GA from entering the local optimal solution.

(3) **Mutation:** Assumed that the dimension of encoded chromosome was s , a random integer will be randomly generated within a specified mutation probability P_m . Using generated integers to obtain the newly numbered genes, some new chromosomes were constructed to replace the ones before mutation to form a new population. Mutated individuals in the population can maintain the diversity of the population, and can also prevent the occurrence of premature. Similarly, a same judgment criterion was performed on the iterations to prevent newly generated chromosomes from being the same as the parent.

Generally, the individual-specific MI-related features were come from different time segment and frequency bands. The introduced GA based TSFBCSP features selection encoded such features to genes, and utilized the randomness of GA to avoid the selection process falling into local optimal solution. Based on the characteristics of GA, the feature selection process can be extended to the parallel framework to improve the efficiency.

3.2. Decoding TSFBCSP features for fitness function

Support vector machine (SVM) [54] has a good performance in limited samples, nonlinear relationships, and multi-categories classification. The proposed TSFBCSP features reserved the spatial-frequency-sequential relationships, which were suitable for the classification using SVM. The SVM classifier constructs the optimal hyper-plane between different categories, and makes a decision to distinguish different categories by maximizing extent. The fitness function of GA for selecting TSFBCSP features was formulated by the average cross-validation classification accuracy of SVM. For solving the nonlinear classification problems, the SVM generally adopts the kernel function to transform nonlinear problems into linear problems. Based on the assumption of two categories, for the l extracted TSFBCSP features and the corresponding MI tasks, the classification problem of MI can be represented as $(x_i, y_i), i = 1, 2, \dots, l, x \in R^N$, where represent the two different categories of MI. The decision function of SVM can be expressed as:

$$y_i[(w \cdot x_i) + b] - 1 \geq 0, i = 1, 2, \dots, l \quad (8)$$

According to the solver of SVM [54], the above problem can be expressed to the optimization objective function:

$$\begin{aligned} \min Q(a) &= \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j \cdot K(x_i, x_j) - \sum_{i=1}^l a_i \\ \text{s.t. } &\sum_{i=1}^l a_i y_i = 0, 0 \leq a_i \leq C \end{aligned} \quad (9)$$

where C represents the penalty factor for the features. We retained the sequential relationships in TSBFCSP features, so the polynomial kernel function was selected to better express the sequential relationships:

$$K(x, x_i) = (\gamma x^T x_i + r)^d \quad (10)$$

where γ is the parameter for the polynomial function. Finally, the optimal hyper-plane for classifying MI tasks can be expressed as:

$$f(x) = \operatorname{sgn} \left(\sum_{i=1}^l a_i^* y_i K(x, x_i) + b^* \right) \quad (11)$$

where a_i^* and b^* represent the optimal decision hyper-plane, which is provided by the support vectors. In the experiment, by using the libSVM

model [55], the one vs rest strategy was also applied to meet the process of TSFBCSP features extraction.

3.3. MapReduce based parallel genetic algorithm

In fact, the fitness function computations of the aforementioned TSFBCSP-GA feature selection have a relatively high computational complexity. The features decoding and selection on all combinations of time segments and frequency bands are the same. To improve the efficiency of TSFBCSP-GA features selection, we adopted the extension of parallel GA based on MapReduce framework (TSFBCSP-MRPGA) [50]. The MapReduce is a parallel computing model that specializes in processing big data, and the data is stored in shards on cluster computing nodes. The entire computation process based on the MapReduce framework can be decomposed into two stages: Map stage and Reduce stage. In the Map stage, each computing node on cluster starts Map tasks. The Map task reads data shards in parallel, and decomposes each row of data in the shards into key-value pairs($key_1, value_1$). The term of key_1 is line number by default setting, and the term of $value_1$ is the actual content data of each row. In this paper, the logic in the process of GA was transformed into new key-value pairs($key_2, value_2$). Then, the processing logic of Map tasks were expressed as follow:

$$\text{Map} :: (key_1, value_1) \rightarrow (key_2, value_2) \quad (12)$$

When the key-value-lists ($key_2, list(value_2)$) were received, the Reduce tasks were processed in accordance with specific logic, and the results of MRPGA were output as($key_3, value_3$). The logic was represented as:

$$\text{Reduce} :: (key_2, list(value_2)) \rightarrow (key_3, value_3) \quad (13)$$

Since the computation of fitness function needs to go through the training phase, it has a large amount of computations in fitness function. Therefore, the computation of fitness function was put into Map stage, and the obtained fitness values were transferred to Reduce stage. During the Reduce stage, the three operators of selection, crossover, and mutation were performed to generate new chromosomes. The population and parameters of the MR-PGA algorithm were initialized in the main function, and the MapReduce processes were executed on multiple clusters to obtain the final MI pattern recognition results. We illustrate the diagram of MRPGA algorithm in Fig. 2.

The MRPGA algorithm for the optimal time segments and frequency bands selected from TSFBCSP features is illustrated in **Algorithm 3**.

Input:	Algorithm 3: The MRPGA algorithm
	TSFBCSP features
Output:	Selected optimal time segments and frequency bands
Steps:	<ul style="list-style-type: none"> (1) Building the Mapreduce framework based GA; (2) Setting the parameters of C and γ for SVM, and the parameters g, k, s, P_c, P_m of for GA;

(continued on next column)

(continued)

Input:	Algorithm 3: The MRPGA algorithm
	TSFBCSP features

- (3) Encode the number of TSFBCSP features as the chromosome of the GA;
- (4) For each slave node on the MapReduce framework, incorporating the chromosome represented number of TSFBCSP features into the SVM classifier, and compute the classification performance as the fitness function;
- (5) For the master node on the MapReduce framework, collect the fitness function computed from each slave node, perform selection operation by Eq. (7), and crossover operation, and mutation operation;
- (6) By exchange chromosome and fitness function among master and slaves after iterations, the optimal time segments and frequency bands were selected by the best fitness function;
- (7) **return** Optimal time segments and frequency bands

In total, we combination the Algorithm 1, 2, and 3 to form the proposed TSFBCSP-MRPGA algorithm, and the algorithm is summarized in **Algorithm 4**.

Input:	Algorithm 4: The TSFBCSP-MRPGA algorithm
	Multivariate EEG signals dataset

Output:	The optimal selected TSFBCSP features and the best classification accuracy
----------------	--

- Steps:**
 - (1) Extract EEG samples set from **Algorithm 1**;
 - (2) Compute TSFBCSP features from **Algorithm 2**;
 - (3) Select the optimal time segments and frequency bands by **Algorithm 3**;
 - (4) Save the selected features and the best classification accuracy;
 - (9) **return** optimal selected features and best classification accuracy

Fig. 3 illustrates the overall framework for the proposed TSFBCSP-MRPGA algorithm. During the training phase, the MapReduce framework will help us to speed up the TSFBCSP features selection procedure, and the optimal TSFBSCP features are selected for each subject. After training phase, the optimal subject-specific and session-specific TSFBCSP features were selected for the testing phase. Finally, the optimal specific features will help to improve the classification performance of MI tasks.

4. Experiments and analysis

4.1. Experimental setups

To validate the feasibility and effectiveness of the proposed TSFBCSP-GA/MRPGA algorithm, two public benchmark datasets from BCI Competition IV [56], dataset 2a and dataset 2b, were selected for experiments. Both datasets contained MI-EEG signals collected from 9 different subjects. There are four MI tasks for dataset 2a (left, right, foot, and tongue) and two MI tasks for dataset 2b (left and right). Both datasets have the same experimental paradigm of performing MI tasks,

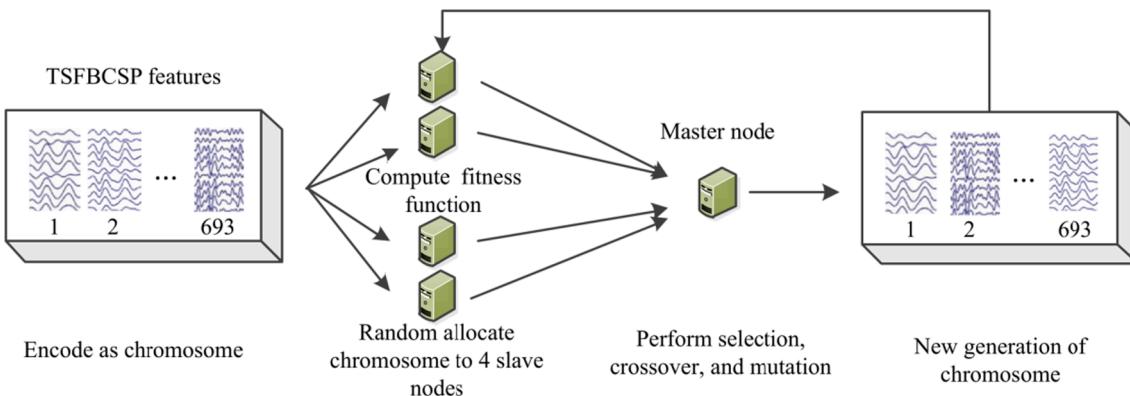


Fig. 2. The diagram of MRPGA algorithm for optimal features selection.

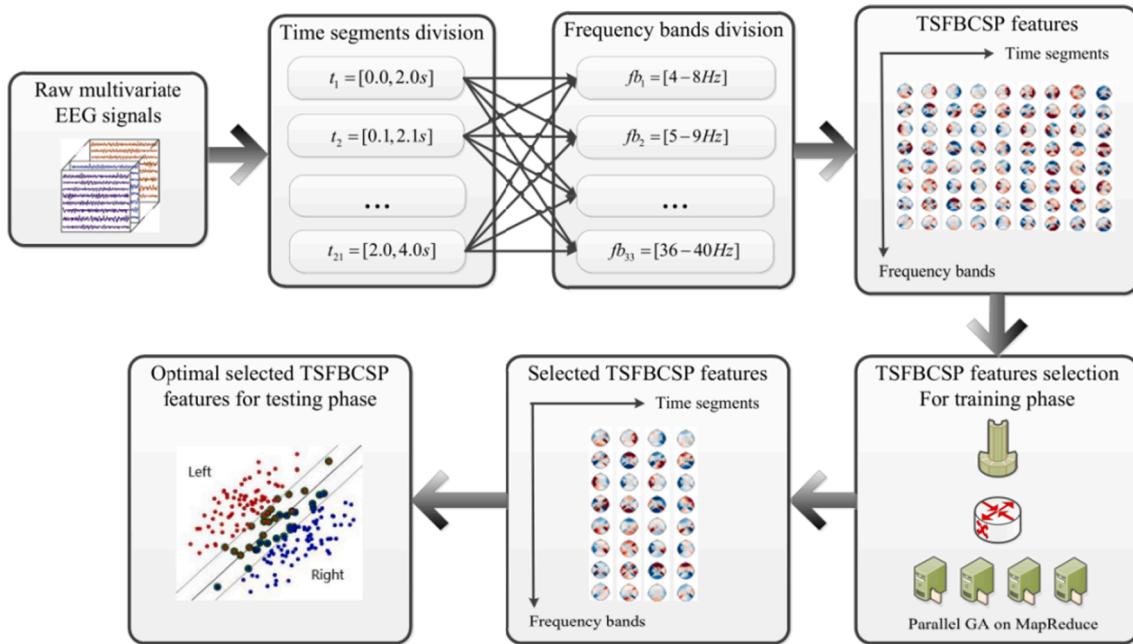


Fig. 3. The overall framework for the proposed TSFBCSP-MRGPA algorithm.

as shown in Fig. 4. At beginning, a warning beep and a cross on the screen were presented to the subject to keep the focus. Then, a cue of arrow is presented on the screen to guide the class of MI tasks for 1.25 s. The subject began to perform MI task based on the guidance of the cue after it presented 1 s, and each MI task continue 4 s. At the end of performing MI task, a randomly time duration was given to the subject for resting.

There were two recording sessions for dataset 2a, and each session contained 288 labeled EEG trials for four classes. The EEG recording device contained 22 electrodes. For the pattern recognition experiment, the 288 EEG trials from the first session were used for training, and the 288 EEG trials for the second session were used for testing. There are five collecting sessions in dataset 2b. The first two sessions contained 120 labeled EEG trials, and the last three sessions contained 160 labeled EEG trials, with a total of 720 EEG trials for two classes. The EEG recording device contained 3 electrodes. For the pattern recognition experiment, the 400 EEG trials from the first three recording sessions were used for training, and the 320 EEG trials from the last two recording sessions were used for testing.

To conduct a fair comparative experiment, our software environment was Matlab R2018b in Windows 10 operating system. The platform of performing MRPGA was Hadoop 3.1.3, JDK 1.8.0–162, Eclipse 4.7.0, and Ubuntu Kylin 16.04. The Hadoop cluster for experiment contained 5 nodes, which were connected in a master–slave manner. The hardware environment of the master node was Intel(R) Core(TM) i9-10885H CPU @ 2.40 GHz with 32 GB memory, and the four slave nodes was Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz with 16 GB memory. The

master–slave environment was connected with the internet. The crossover and mutation operations were dispensed to the slave nodes, and the obtained chromosomes were uploaded to the master node to compute the fitness function of the selection phase.

To compare the classification performance and efficiency of the proposed algorithm, the FBCSP + SVM algorithm was deemed to the baseline. For the comparison of feature-level, the fwr-csp [57], Bayesian optimization [58], and time-contained spatial filtering [59] were introduced to compare the classification performance in the feature-level. Note that, the reference of TSF features were classified using a deep learning model (STAN), and we exhibited the classification results using the SVM classifier. Ideally, the proposed TSFBCSP-GA architecture can also be extended to the deep learning model, whilst it takes a lot of time to train such model. To give the ablation study of feature extraction and selection, the selected FBCSP features by GA, and the entity of TSFBCSP features were compared in the feature-level. For the comparison of model-level, five recent deep learning-based models, ConvNet [60], EEGNet [61], MCNN [62], C2CM [63], DCNN [64] were introduced to be compared with the proposed algorithm.

4.2. Comparison results of motor imagery classification and efficiency

In the actual comparative experiments, the training set of both datasets was decomposed into ten equal parts, and a 10×10 leave one cross-validation were used to train a robust SVM model on the training sets. The well-trained SVM model then incorporated into testing set to obtain the average classification accuracy, and such accuracy was utilized to evaluate the fitness of GA/MRPGA. During the experiment, due to the constraint of computing resources, we reduced the sampling rate of multivariate EEG signals to 125 Hz. Table 1 shows the dataset parameters settings for the comparative experiments.

To show the experimental results, Table 2 and Table 4 showed the experimental results of feature-level comparisons on dataset 2a and 2b, respectively. Table 3 and Table 5 showed the experimental results of model-level comparisons on dataset 2a and 2b, respectively. In these tables, the “Proposed” means TSFBCSP-GA and the “Proposed-MR” means TSFBCSP-MRPGA, and the “SD” in the tables means the standard deviation among all subjects. Among them, the parameter of CSP algorithm was set to $m = 1$ for TSFBCSP features on both datasets, and the

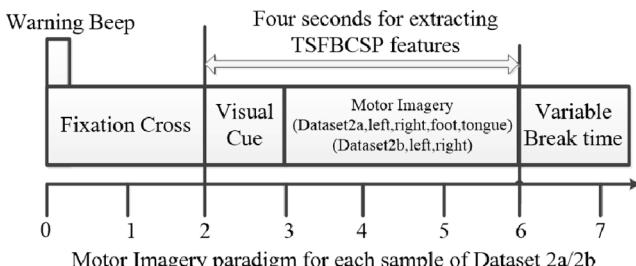


Fig. 4. The MI paradigm of each trial for the selected two datasets.

Table 1

The basic information, time segments, and frequency bands of dataset 2a and dataset 2b.

Dataset 2a			Dataset 2b		
Training set	Validating set	Testing set	Training set	Validating set	Testing set
260	28	288	360	40	320
Electrodes	Sampling rate	Duration	Electrodes	Sampling rate	Duration
22	125 Hz	4000 ms	3	125 Hz	4000 ms
Time segments (points)	[1, 250], [25, 275], ..., [250, 500]		Time segments (points)	[1–250], [25, 275], ..., [250, 500]	
Frequency bands (Hz)	[4,8], [5,9], ..., [36,40]		Frequency bands (Hz)	[4,8], [5,9], ..., [36,40]	

parameters of SVM was set to $C = 0.5, \gamma = 1$ on Dataset 2a and $C = 0.01, \gamma = 0.01$ on dataset 2b. The default parameters of GA/MRPGA: the number of chromosomes was $g = 30$, the dimension of each chromosome was $s = 20$, the probability of crossover was $P_c = 95\%$, and the probability of mutation was $P_m = 5\%$. The compared algorithms used the same setting and the CNN model using the default setting from the references, and extended to the same classification tasks, environment,

and strategies.

From the results in Table 2, it can be seen that the proposed TSFBCSP-GA/MRPGA has the best recognition performance on dataset 2a, which are 74.54% and 74.00%, respectively. At the same time in Table 4, the proposed TSFBCSP-GA/MRPGA has the best recognition performance on dataset 2b, which are 79.78% and 79.29%, respectively. In addition, we have found that the entity of TSFBCSP achieves a higher classification performance than the baseline. Due to the prevention of redundancy parts of TSFBCSP, the selected TSFBCSP by GA and MRPGA is capable of achieving better classification performance. However, the GA performed on the conventional FBCSP obtains a comparable classification performance than FSDE and FSV, but less than the GA and MRPGA performed on the proposed TSFBCSP. This ablation study has shown that the proposed TSFBCSP-GA/MRPGA algorithm achieves a large classification performance improvement on both datasets. The classification performance of the proposed algorithm is significantly improved than the baseline using a statistical student *t*-test ($p < 0.05$) on both datasets.

From the results in Table 3 and Table 5, we have found that the proposed TFCSP-MRPGA outperforms the conventional deep learning models, ShallowNet and EEGNet. For the recent deep learning models, the MCNN model and DCNN model achieved a little bit higher

Table 2

Comparison classification results for the feature-level of dataset 2a.

	Baseline	FBCSP-GA	TSFBCSP	FWR-CSP	TSF-SVM	Proposed	Proposed-MR
S1	77.4	83.33	85.76	72.4	72.5	86.11	86.46
S2	54.2	54.86	57.64	65.0	76.7	61.81	59.03
S3	69.8	86.11	88.89	78.2	82.1	86.81	89.24
S4	56.3	63.89	66.67	61.5	68.7	71.18	69.44
S5	46.9	57.29	61.46	68.6	52.3	62.85	63.19
S6	52.1	50.69	54.17	55.9	55.0	56.25	54.51
S7	83.0	82.29	86.81	76.3	86.5	90.28	87.15
S8	60.4	75.35	78.47	80.8	87.3	80.56	80.21
S9	68.4	77.43	80.21	68.3	83.2	78.47	81.60
Mean	63.17	70.14	73.34	69.67	73.81	74.92	74.54
SD	12.19	13.57	13.47	8.13	13.00	12.39	13.21

Table 3

Comparison classification results for the model-level of dataset 2a.

	ConvNet	EEGNet	MCNN	C2CM	DCNN	Proposed	Proposed-MR
S1	86.6	85.0	90.2	87.5	80.3	86.11	86.46
S2	62.3	56.6	63.4	65.3	75.8	61.81	59.03
S3	89.9	81.7	89.4	90.3	74.4	86.81	89.24
S4	65.6	66.4	71.2	66.7	58.3	71.18	69.44
S5	55.2	54.9	62.8	62.5	76.8	62.85	63.19
S6	48.5	59.6	47.7	45.5	63.1	56.25	54.51
S7	86.1	92.3	90.1	89.6	78.4	90.28	87.15
S8	78.4	75.7	83.7	83.3	83.3	80.56	80.21
S9	76.1	74.8	82.3	79.5	78.2	78.47	81.60
Mean	72.10	71.89	75.64	74.47	74.29	74.92	74.54
SD	14.83	13.28	15.14	15.33	8.21	12.39	13.21

Table 4

Comparison classification results for the feature-level of dataset 2b.

	Baseline	FBCSP-GA	TSFBCSP	FWR-CSP	BO	Proposed	Proposed-MR
S1	70.3	70.00	72.19	77.2	71.4	75.63	72.19
S2	55.4	59.29	61.07	59.3	60.6	62.50	61.43
S3	55.6	59.69	59.69	62.4	58.1	61.25	61.56
S4	94.7	96.88	96.88	94.8	97.1	97.81	97.50
S5	80.6	84.38	84.69	85.7	91.3	87.81	85.63
S6	80.0	87.19	85.31	84.6	86.0	85.00	87.19
S7	74.1	79.69	79.69	77.4	77.0	81.25	80.31
S8	79.7	90.00	89.38	84.6	92.0	92.81	90.00
S9	76.3	80.31	80.31	79.2	84.5	85.31	82.19
Mean	74.08	78.60	78.80	78.36	79.78	81.04	79.78
SD	12.47	13.13	12.46	11.31	13.96	75.63	72.19

Table 5

Comparison classification results for the model-level of dataset 2b.

	ConvNet	EEGNet	MCNN	C2CM	DCNN	Proposed	Proposed-MR
S1	72.0	73.8	74.5	74.8	76.8	75.63	72.19
S2	57.0	56.7	62.6	61.3	61.2	62.50	61.43
S3	64.9	64.5	65.3	65.5	67.5	61.25	61.56
S4	94.4	93.2	98.8	94.4	96.8	97.81	97.50
S5	89.9	81.9	86.7	86.7	88.2	87.81	85.63
S6	83.3	85.8	88.6	87.5	87.7	85.00	87.19
S7	78.1	72.7	80.5	79.4	79.6	81.25	80.31
S8	90.8	91.5	91.8	89.6	92.7	92.81	90.00
S9	77.9	72.5	81.5	81.7	82.6	85.31	82.19
Mean	78.70	76.96	81.14	80.10	81.46	81.04	79.78
SD	12.48	12.20	12.00	11.13	11.61	75.63	72.19

classification performance than the proposed algorithm. Since more complex network architectures introduced in the recent deep learning models, there are more computational quantities and regularizations in such models, resulting in a larger computational complexity and more complex un-interpretability. As can be seen in **Table 6**, the proposed TSFBCSP-MRPGA algorithm consumes a similar computational time with the classical ShallowNet and EEGNet, which is more efficient than the recent MCNN, C2CM, and DCNN deep learning models. Without the requirement of GPUs, the proposed algorithm can be employed to the embedded BCI applications with limited environment.

The results of each iteration for the GA feature selection process were shown in **Fig. 5**. From the results in **Fig. 5**, it can be seen that utilizing GA/MRPGA for the selection of TSFBCSP features has high robustness. As the iterations progress, the average recognition accuracies of different subjects have achieved different degrees of improvement. In the setting of GA/MRPGA feature selection, if there was a better recognition accuracy on a certain time segment and frequency band, they will be retained in the next generation of iteration. For the time segment and frequency band have negative impact on recognition accuracy, they will be replaced to new chromosomes in the next generation of iteration. After several rounds of iterations, the optimal time segments and their corresponding frequency bands were retained for the testing phase. Therefore, the proposed TSFBCSP-GA/MRPGA can almost surpass the conventional compared algorithms.

In the case where the TSFBCSP-GA and TSFBCSP-MRPGA obtained the best average recognition accuracy, **Fig. 6** and **Fig. 7** gave the optimal time segment and frequency bands on the numbers of selected TSFBCSP features for both datasets. Among them, the results of Subject 3 in Dataset 2a were displayed, and the results of Subject 8 in Dataset 2b were displayed. From the results in **Fig. 6** and **Fig. 7**, it can be seen that although the EEG signal has non-linear and non-stationary characteristics, the proposed GA/MRPGA can still select the most representative TSFBCSP feature, and finally improve the average recognition accuracy. Compared with **Fig. 6** (a) and (b) and **Fig. 7** (a) and (b), we have found that the effective TSFBCSP features selected by classical GA (blue squares) and MRPGA (orange squares) were almost the same. These comparative results have proved that the parallel GA based on MapReduce was effective, and the near-optimal TSFBCSP features were selected by the MRPGA for both datasets. Meanwhile, the efficiency of

feature selection was greatly improved by parallel processing. Moreover, although a large number of redundant features were extracted by a detailed decomposition on time domain and frequency domain, judging from the distributions of blue squares and orange squares, the truly effective features for specific subject and were few and sparse. The testing phase has no more obsessions on the efficiency of redundant features.

In addition, **Table 6** showed the time complexity of several comparative algorithms. In both datasets, the baseline only needs to extract classic FBCSP features, so the time complexity of feature extraction and training is the lowest. In the training phase, the FSDE and FSV strategies performed regularizations on features during recognition, so the training process contained more time-consuming. When directly performing GA on feature selection, a large time consumption required for several iterations, so the training time complexity was highest on both datasets. To reduce the time complexity, the parallel GA was constructed on a MapReduce framework with 5 nodes. The time training time complexity of MRPGA was obviously reduced, reaching the same order of magnitude as ShallowNet and EEGNet. For the testing phase, the TSFBCSP features only need to extract on the optimal selected numbers on specific time segments and frequency bands, so the time complexity was not much different from the classical baseline algorithm. Nevertheless, ShallowNet and EEGNet belonged to the end-to-end model and such deep learning models no longer distinguish between feature extraction and classification, resulting in a low time complexity in the testing phase, too. In all, the proposed TSFBCSP-MRPGA took the advantages of both the parallel capability of MapReduce framework and the supervised feature selection capability of GA. It achieved a good classification performance while also having an acceptable time complexity.

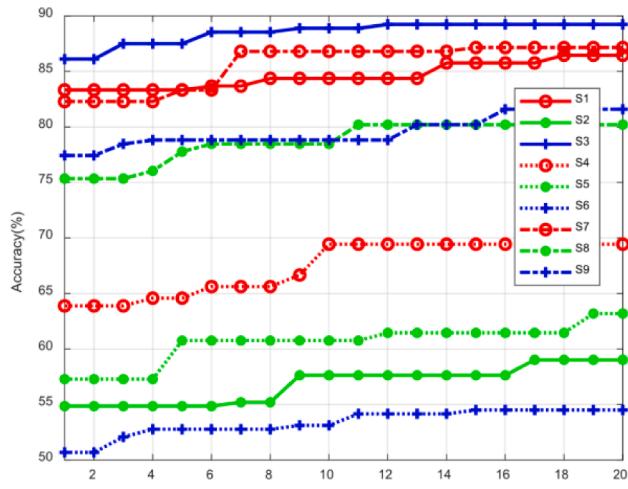
4.3. Ablation study

To further analyze the recognition performance of the proposed TSFBCSP-GA/MRPGA, an ablation study was performed on the number of chromosomes, the length of chromosomes, and the visualization of selected TSFBCSP features.

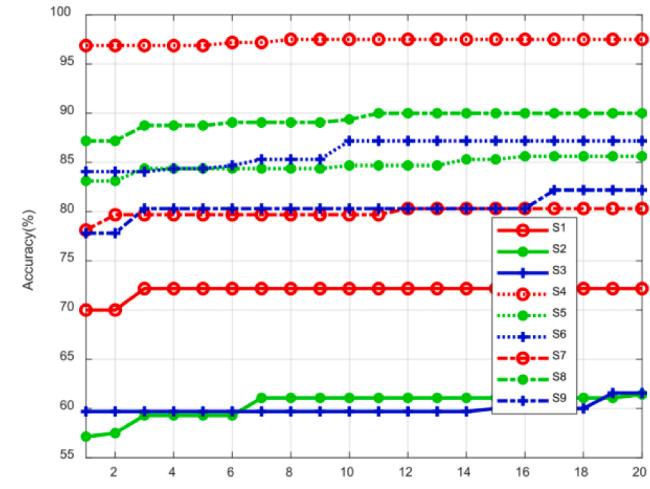
Table 6

Time complexity comparison for the comparative algorithms.

Average computational time complexity	Dataset2a			Dataset2b		
	Feature Extraction	Training Model	Testing Model	Feature Extraction	Training Model	Testing Model
Baseline	14.75 s	11.51 s	15.33 s	22.46 s	9.60 s	9.60 s
FWR-CSP	16.82 s	28.62 s	16.81 s	24.62 s	10.53 s	10.62 s
TSF	18.82 s	–	–	26.46 s	–	–
ShallowNet (GPU)	–	59.3 min	4.45 s	–	90.1 min	6.76 s
EEGNet (GPU)	–	22.9 min	4.72 s	–	34.8 min	6.61 s
TSFBCSP-GA	17.75 s	224.0 min	13.62 s	28.10 s	217 min	11.23 s
TSFBCSP-MRPGA	17.75 s	44.8 min	13.62 s	28.10 s	43.4 min	11.23 s



(a) Dataset 2a



(b) Dataset 2b

Fig. 5. Each iteration of TSFBCSP-GA on both datasets for each subject.

4.3.1. Number of chromosomes for GA/MRPGA

Due to the limit of hardware performance, we selected a range of chromosomes numbers $g = [10, 20, 30, 40, 50, 60, 70]$ for testing. Based on the selected numbers of chromosomes, we have verified the performance of TSFBCSP-GA/MRPGA on both datasets. Table 7-10 illustrated the average recognition accuracy for the different number of chromosomes. From the results in Table 7 and 8 on Dataset 2a, it can be seen that the GA/MRPGA have the best performance between 30 and 50 chromosomes. Among them, the best average recognition accuracy (74.65%) of GA achieved at chromosomes number of 50, and the best average recognition accuracy (74.23%) of MRPGA achieved at chromosomes number of 50, too. From the results in Table 9 and 10 on Dataset 2b, it can be seen that the GA/MRPGA have the best performance between 50 and 70 chromosomes. Among them, the best average recognition accuracy (81.04%) of GA achieved at chromosomes number of 60, and the best average recognition accuracy (80.58%) of MRPGA achieved at chromosomes number of 70. Compared with GA, although MRPGA has slight performance degradation in TSFBCSP features selection, MapReduce framework will bring extremely parallel efficiency improvement in training phase.

From the aforementioned and analyzed results, we have suggested that the optimal number of chromosomes required by GA/MRPGA was almost the same, but the range of the optimal number of chromosomes was determined on the distribution of EEG signals in different datasets. Within the range of the optimal number of chromosomes, how to select the optimal number was a compromise choice: on the one hand, a smaller number of chromosomes will cause recognition performance degradation; on the other hand, a larger number of chromosomes will further increase the computational time complexity. When the increase of chromosomes does not bring about a large performance improvement, generally the smallest number in the suitable range can be utilized for TSFBCSP-GA/MRPGA feature selection, and applied for actual BCI applications.

4.3.2. Length of genes on chromosomes for GA/MRPGA

According to the results in Table 5-8, the number of chromosomes to obtain the best average recognition accuracy was respectively selected for both datasets, Dataset 2a ($g = 30$) and Dataset 2b ($g = 60$), to verify the impact of different lengths of encoding genes on chromosomes for GA/MRPGA. Due to the limit of hardware performance, we selected a range of chromosomes numbers $s = [10, 20, 30, 40, 50, 60, 70]$ for testing. Based on the selected lengths of chromosomes, we have verified the performance of TSFBCSP-GA/MRPGA on both datasets. Table 11-14

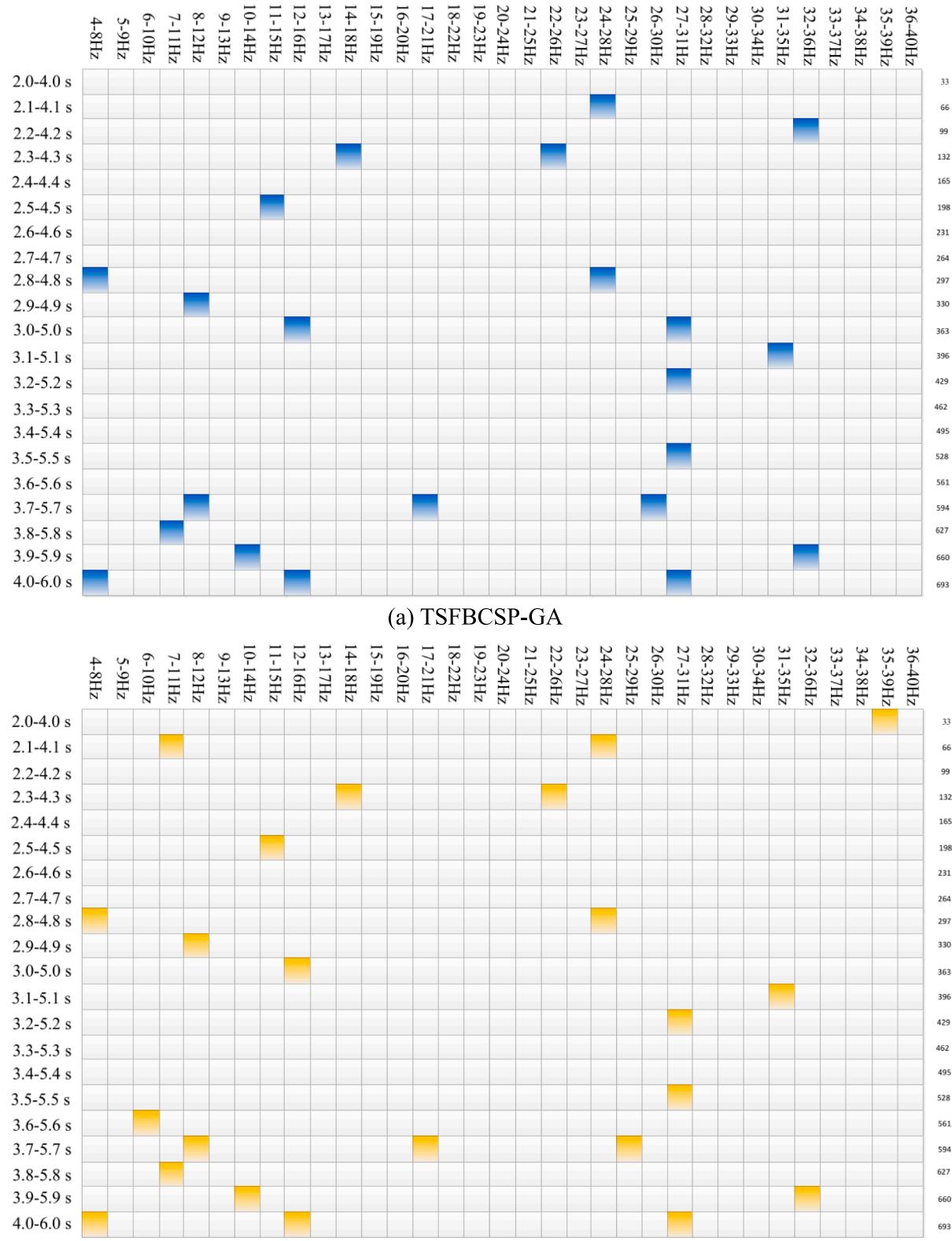
illustrated the average recognition accuracy for the different length of genes for chromosomes. From the results in Table 11 and 12 on Dataset 2a, it can be seen that the best recognition performance of GA (75.35%) was based on a length of 20 genes, while the best recognition performance of MRPGA (74.89%) was based on a length of 10 genes. From the results in Table 13 and 14 on Dataset 2b, it can be seen that the best recognition performance of GA (80.97%) was based on a length of 10 genes, while the best recognition performance of MRPGA (80.44%) was based on a length of 10 genes, too.

Similarly, MRPGA has a certain recognition performance degradation due to the parallel processing, and the degradation on different lengths of genes on chromosomes is not large. At the same time, the optimal gene lengths of chromosomes were almost the same between GA and MRPGA. Due a wide range of influence for the number of chromosomes on the optimal recognition performance, the TSFBCSP-GA/MRPGA should be compromised between performance attenuation and time complexity increase. Different from the length of chromosomes, the length of encoding genes on chromosomes has a limited influence for achieving the optimal recognition performance. According to the TSFBCSP features extracted from different EEG datasets, a stable length of encoding genes on chromosomes can be utilized for TSFBCSP-GA/MRPGA feature selection, and applied for actual BCI applications.

4.3.3. Visualization of selected TSFBCSP features

For the best TSFBCSP features selected in Fig. 6 and Fig. 7, a dimension reduction visualization tool, t-SNE [65], was used for feature visualization. Fig. 8 and Fig. 9 respectively showed the feature visualization of different categories under the optimal selected TSFBCSP features on subject-specific time segments and frequency band combination in both datasets. For horizontal comparison, we also showed the results of classical FBCSP features in the baseline algorithm. For the visualization, all CSP features from each EEG trial were reduced to two dimensions through t-SNE and displayed as scatters. Among them, the four MI categories in Dataset 2a were displayed as left hand (blue), right hand (green), feet (red), and tongue (cyan), and left hand (blue) and right hand (red) were displayed from Dataset 2b.

According to the results in Fig. 8 and Fig. 9, it shows that no matter for the Dataset 2a or Dataset 2b, the TSFBCSP features selected by the GA/MRPGA was better than the classical FBCSP features. Whether it was Dataset 2a of four MI categories or Dataset 2b of two MI categories, the selected TSFBCSP features emerged better separability in the visualization of t-SNE, while the discriminant of FBCSP features was weaker than the selected TSFBCSP features. Therefore, in the actual MI-BCI

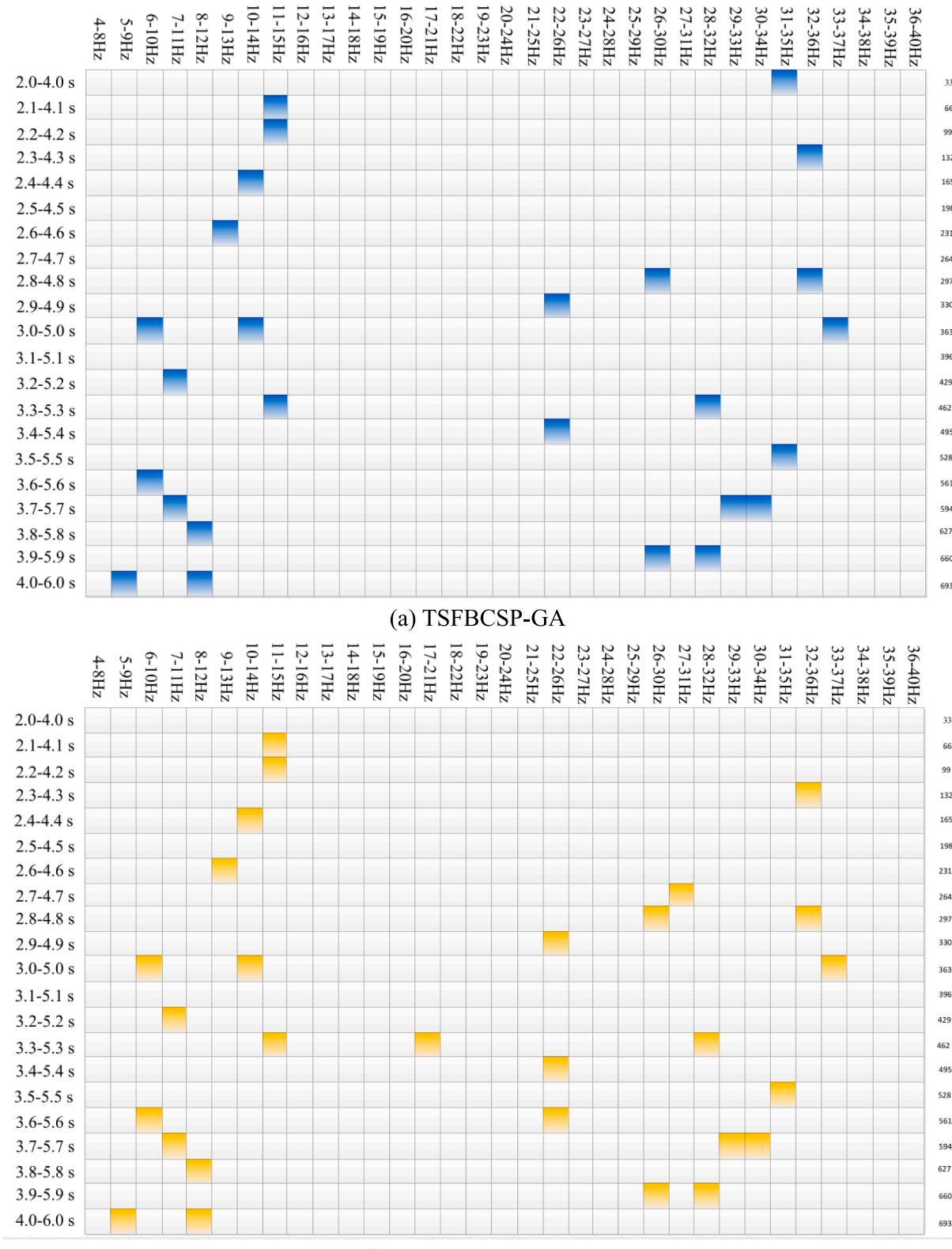


(c) TSPBCSI-MRI GA

applications, we need to first select a suitable TSBFCSP features according to the training phase of GA/MRPGA, and then applied the

Different from the classical FBCSP features, the optimal TSBFCSP-GA selected TSFBCSP features to the BCI control devices.

lasted for a long time, and the subjective imagination of different subjects may be varied for different parts of body in different time periods. Therefore, the simple decomposition of time segment and frequency band will lose the vital MI-related features, resulting in a decrease for the average accuracy of pattern recognition. Therefore, the



decompositions of detailed time segments and frequency bands can obtain more detailed brain features variations in each MI category, and thus have an excellent improvement in the average accuracy of pattern recognition. Meanwhile, the parallel MapReduce framework was used to construct PCA, and the optimal TSEBCSP features could be selected.

- 7 -

This study has investigated a detailed TSEPCSP features extraction

Table 7

Average classification accuracy under different number of chromosomes using TSFBCSP-GA algorithm for dataset 2a.

Number of Chromosome	$g = 10$	$g = 20$	$g = 30$	$g = 40$	$g = 50$	$g = 60$	$g = 70$
S1	84.36	82.29	86.46	85.42	87.50	83.68	84.72
S2	60.42	59.03	59.03	60.07	57.29	60.42	55.56
S3	84.72	87.50	89.24	88.89	90.28	86.11	85.76
S4	71.88	63.89	69.44	69.10	69.10	66.32	64.58
S5	55.21	60.76	63.19	61.46	62.50	57.29	59.03
S6	48.96	51.73	54.51	56.60	59.03	53.82	51.74
S7	83.33	86.81	87.15	85.42	85.42	88.19	89.58
S8	74.65	76.74	80.21	79.17	79.86	81.25	78.47
S9	77.78	82.64	81.60	79.86	80.90	76.74	78.47
Mean	71.26	72.38	74.54	74.00	74.65	72.65	71.99

Table 8

Average classification accuracy under different number of chromosomes using TSFBCSP-MRPGA algorithm for dataset 2a.

Number of Chromosome	$g = 10$	$g = 20$	$g = 30$	$g = 40$	$g = 50$	$g = 60$	$g = 70$
S1	84.36	80.90	86.46	85.42	87.15	83.68	84.72
S2	60.42	57.64	57.64	58.68	56.94	60.07	55.56
S3	82.97	87.15	89.24	88.89	89.93	86.11	85.76
S4	71.88	63.89	69.44	68.75	68.40	66.32	63.19
S5	54.86	57.64	61.46	61.11	62.50	57.29	58.33
S6	48.96	51.73	54.17	55.21	56.94	53.13	51.74
S7	82.29	86.46	87.15	85.42	85.42	87.85	89.24
S8	72.22	76.39	80.21	78.82	79.86	80.90	78.47
S9	77.08	81.94	80.21	79.51	80.90	76.74	78.13
Mean	70.56	71.53	74.00	73.53	74.23	72.45	71.68

Table 9

Average classification accuracy under different number of chromosomes using TSFBCSP-GA algorithm for dataset 2b.

Number of Chromosome	$g = 10$	$g = 20$	$g = 30$	$g = 40$	$g = 50$	$g = 60$	$g = 70$
S1	65.63	72.81	72.19	75.00	73.75	75.63	74.69
S2	57.86	60.00	61.43	61.07	61.07	62.50	65.36
S3	57.81	60.94	61.56	59.06	59.06	61.25	60.63
S4	96.56	97.50	97.50	98.13	97.50	97.81	97.81
S5	82.81	86.25	85.63	88.75	86.88	87.81	87.50
S6	83.13	81.25	87.19	83.75	86.56	85.00	84.38
S7	75.31	81.25	80.31	80.94	81.56	81.25	82.19
S8	88.75	90.31	90.00	91.56	91.88	92.81	90.94
S9	79.69	80.63	82.19	85.00	84.69	85.31	85.31
Mean	76.39	78.99	79.78	80.36	80.33	81.04	80.98

Table 10

Average classification accuracy under different number of chromosomes using TSFBCSP-MRPGA algorithm for dataset 2b.

Number of Chromosome	$g = 10$	$g = 20$	$g = 30$	$g = 40$	$g = 50$	$g = 60$	$g = 70$
S1	65.63	72.81	72.19	72.50	73.75	75.00	74.38
S2	57.86	60.00	61.07	60.00	61.07	61.79	64.64
S3	57.81	59.69	60.00	59.06	59.06	61.25	60.63
S4	96.56	96.88	97.19	97.81	97.19	97.19	97.50
S5	82.81	85.63	85.31	88.44	86.25	86.56	87.50
S6	83.13	81.25	87.19	83.44	85.94	84.38	84.06
S7	75.31	79.69	80.31	80.94	80.94	80.63	81.56
S8	87.50	90.31	90.00	91.56	91.56	91.56	90.63
S9	79.38	80.63	80.31	83.13	83.75	84.38	84.31
Mean	76.22	78.54	79.29	79.65	79.95	80.30	80.58

from EEG signals, and a MapReduce based parallel GA was used to select the discriminative features for more efficiency and better classification performance without the requirements of GPUs. The motor imagery

Table 11

Average classification accuracy under different length of genes for chromosomes using TSFBCSP-GA algorithm for dataset 2a.

Length of Genes	$s = 10$	$s = 20$	$s = 30$	$s = 40$	$s = 50$	$s = 60$	$s = 70$
S1	86.46	85.07	86.46	85.42	84.72	85.42	86.11
S2	59.03	60.42	61.11	61.81	61.81	61.11	61.81
S3	89.24	89.24	87.85	86.46	87.50	87.85	86.81
S4	69.44	71.88	69.44	69.10	66.32	68.75	71.18
S5	63.19	66.32	61.11	62.50	60.42	59.03	62.85
S6	54.51	55.90	56.60	56.60	57.99	56.25	56.25
S7	87.15	89.24	90.28	88.89	90.63	90.63	90.28
S8	80.21	81.25	79.51	80.56	82.29	79.86	80.56
S9	81.60	78.82	79.51	79.51	79.86	80.56	78.47
Mean	74.54	75.35	74.65	74.54	74.62	74.38	74.92

Table 12

Average classification accuracy under different length of genes for chromosomes using TSFBCSP-MRPGA algorithm for dataset 2a.

Length of Chromosome	$s = 10$	$s = 20$	$s = 30$	$s = 40$	$s = 50$	$s = 60$	$s = 70$
S1	86.46	84.72	85.42	85.07	84.38	85.07	85.76
S2	57.64	59.38	61.11	61.11	60.07	61.11	61.11
S3	89.24	88.89	87.85	86.11	87.15	87.50	86.81
S4	69.44	71.18	68.06	69.10	65.97	68.75	70.14
S5	61.46	65.63	61.11	60.76	59.38	59.03	62.85
S6	54.17	55.56	56.60	56.60	56.94	55.21	55.90
S7	87.15	88.54	90.28	88.89	90.28	90.28	90.28
S8	80.21	81.25	79.17	79.86	81.94	79.86	79.51
S9	80.21	78.82	78.13	78.82	79.17	79.17	77.78
Mean	74.00	74.89	74.19	74.04	73.92	74.00	74.46

Table 13

Average classification accuracy under d different length of genes for chromosomes using TSFBCSP-GA algorithm for dataset 2b.

Length of Chromosome	$s = 10$	$s = 20$	$s = 30$	$s = 40$	$s = 50$	$s = 60$	$s = 70$
S1	75.62	73.44	75.31	74.69	74.69	73.75	73.75
S2	62.50	59.64	62.50	62.14	62.14	63.93	63.93
S3	61.25	59.38	59.69	60.63	59.38	59.38	59.38
S4	97.19	97.19	97.50	96.88	97.19	97.19	97.19
S5	87.81	85.00	85.63	85.63	83.44	83.13	83.13
S6	85.00	85.94	87.19	85.31	85.63	86.88	86.88
S7	81.25	79.69	78.13	81.25	80.31	76.88	76.88
S8	92.81	87.81	88.44	88.13	86.56	86.56	86.56
S9	85.31	81.25	82.19	82.50	78.75	81.88	81.88
Mean	80.97	78.82	79.62	79.68	78.68	78.84	78.84

Table 14

Average classification accuracy under different length of genes for chromosomes using TSFBCSP-MRPGA algorithm for dataset 2b.

Length of Chromosome	$s = 10$	$s = 20$	$s = 30$	$s = 40$	$s = 50$	$s = 60$	$s = 70$
S1	75.00	73.44	75.31	73.75	74.38	73.44	73.44
S2	61.79	59.64	61.79	62.14	62.14	62.50	62.50
S3	61.25	59.38	59.06	60.63	59.06	59.06	59.06
S4	97.19	97.19	97.19	96.88	97.19	96.88	96.88
S5	87.81	85.00	85.63	85.63	82.81	82.50	82.50
S6	84.38	85.63	85.63	85.31	85.63	86.25	86.25
S7	80.63	79.38	78.13	81.25	80.00	76.56	76.56
S8	91.56	86.88	88.44	87.81	85.94	86.25	86.25
S9	84.38	80.94	81.25	81.88	78.75	81.56	81.56
Mean	80.44	78.61	79.16	79.48	78.43	78.33	78.33

classification performance of the proposed algorithm transcended the conventional machine learning based algorithms, and achieved better training efficiency compared with the deep learning based algorithms

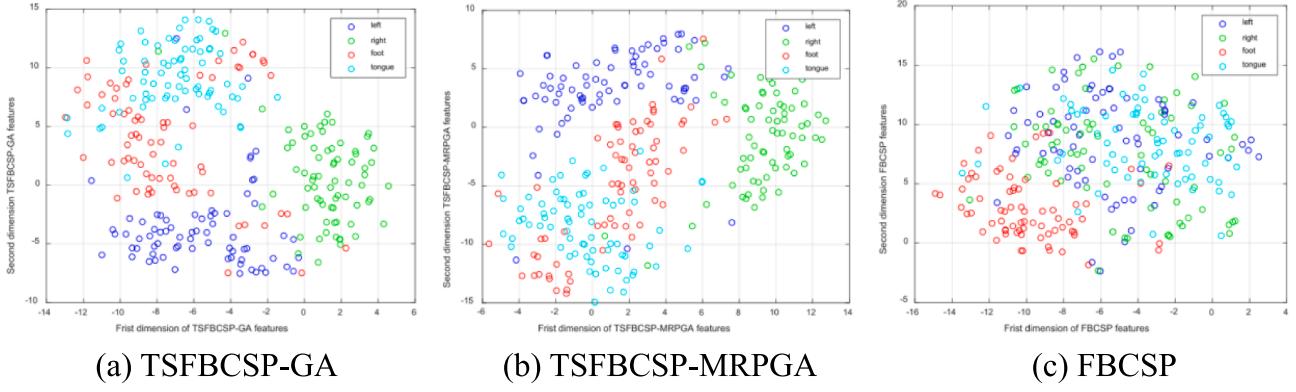


Fig. 8. Optimal time segments with frequency bands features visualization of Dataset 2a by t-SNE.

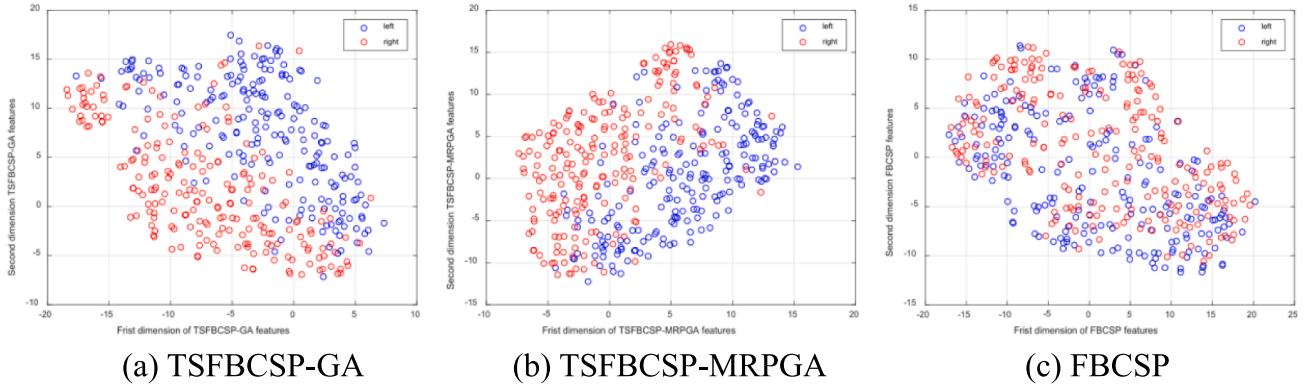


Fig. 9. Optimal time segments with frequency bands features visualization of Dataset 2b by t-SNE.

(see Table 6).

From the results in Fig. 6 and Fig. 7, we have found that the TSFBCSP features selected by the proposed GA/MRPGA algorithm show sparsity on dataset 2a and dataset 2b, and the feature ranges selected by GA or MRPGA algorithm are not much difference. The MapReduce framework improves the iterative efficiency of the genetic algorithm, but retains the classification performance. In fact, this paper considers the minimum intervals of time segments and frequency bands, resulting in a large number of initial candidate features. However, through the selection of the proposed algorithm, the number of features obtained is a little. Among them, the GA/MRPGA algorithm selects 22/23 features from dataset 2a, and selects 26/28 features from dataset 2b. Compared with the total 693 candidate features, the selected features are the most discriminative in time segments and frequency bands, so that achieve the best classification characteristics. At the same time, the selected optimal time segments and frequency bands are related to the specific subjects. If the well-trained classification model is used for online application, the same subjects are used to perform motor imagery tasks. For each sample, the CSP features only need to be extracted on the optimal time segments and frequency bands, which are directly used for motor imagery classification, so as to achieve a high classification performance.

To further prove that the selected features are most discriminative, we used the t-SNE tool for the exhibition of features. The t-SNE tool can minimize the features of any dimensions through manifold and transform them into a 2D space. In the experimental results, we show the results of three different feature, TSFBCSP-GA/ TSFBCSP-MRPGA/ FBCSP, on dataset 2a and dataset 2b (see Fig. 8 and Fig. 9). The classical FBCSP algorithm only considered time segments of [0.5, 4 s] and frequency bands of [4–8] Hz, [8–12] Hz, ..., [36–40] Hz, and the feature distributions of different motor imagery tasks are far less discriminative

than that of TSFBCSP-GA/MRPGA algorithm. That is to say, due to the nonlinear and non-stationary characteristics of motor imagery EEG signals, subjects require different time segments and frequency bands when recording EEG signals across different sessions. Therefore, using directly fixed settings of time segments and frequency bands will loss discriminative features and thus deteriorating classification performance. After feature selection of GA/MRPGA algorithm, the selected features adaptively consider the feature variations across recording sessions, so as to obtain a better classification performance.

There are three advantages of the proposed TSFBCSP-GA/MRPGA algorithm:

(1) **Extendibility:** The experiment showed a comparable classification accuracy and efficiency using four normal nodes in the MapReduce framework (See Table 2-5). Actually, different from the deep learning models, the MapReduce framework has the advantage of extendibility, which improves the classification efficiency by adding a normal computer as a node. During the experiment, we have also compared the classification efficiency of MRPGA using 1/2/3/4 nodes on dataset2a. As shown in Table 15, the classification performance attenuation is not large, but the efficiency is greatly improved. Since the cost of a normal performance computer is very lower than GPU, in real-world MI-BCI applications, the extendibility of the proposed TSFBCSP-MRPGA algorithm is better than deep learning models.

(2) **Practicability:** The proposed TSFBCSP-MRPGA algorithm showed a generalization for MI-EEG signals classification. Compared with the baseline algorithm, the proposed algorithm has a tremendous classification accuracy improvement for all subjects on both datasets. Besides, the ablation studies for different parameters have shown that a suitable number of chromosomes and a suitable length of genes on a chromosome will be generalized to all subjects in a single dataset (See Table 7-14). That is to say, once the parameters of the GA are

determined, it can be applied to the EEG signals of whole subjects collected on a certain acquisition device, and it has excellent practicability in MI classification. In an actual scenario of EEG-based BCI, it is common to acquire a limited number of EEG trials for each subject. Different from the deep learning models that required a large number of training trials, the proposed TSFBCSP-MRPGA algorithm can be trained and applied in a small amount of EEG trials, which can be adopted for many applications such as wearable EEG equipment, and wirelessly transmitted EEG devices.

(3) **Explainability:** Although the proposed TSFBCSP-MRPGA algorithm is slightly inferior to the recent deep learning models, such algorithm is based on the generic classifiers. This paper adopted the SVM classifier, which can be replaced with other classifiers to adapt to the hardware environment, such as LDA, GBDT, etc. The most explainable part is the selected TSFBCSP features by the GA/MRPGA algorithm, the selected TSFBCSP features have more resolutions on the variations of sensorimotor rhythms, which are in accord with the ERD phenomenon of mu and beta rhythms. Based on the “black-box” characteristics of deep learning models, they cannot explain why to obtain a high classification performance and how to select discriminative features. From the explainable BEAMs, we can analyze more detailed variations in the process of MI, and find out the most representative and invariable features through intelligent algorithms, which can be used to build robust and high-performance classifiers.

6. Conclusion

In this paper, we extract TSFBCSP features on more detail decompositions both on time segments and frequency bands. The more detailed decompositions on multivariate EEG signals carry better discriminant information of different MI categories across subjects and recording sessions. To avoid redundant features and improve the efficiency of recognition, we propose a novel parallel GA on MapReduce framework to select the optimal features for recognition. The proposed algorithm is evaluated on two public benchmark EEG signals datasets. The optimal selected TSFBCSP features outperform other variants of CSP features and deep neural networks, but have an acceptable time complexity on pattern recognition.

However, due to the limit computational resources, the proposed GA performed on original CSP features and extended on several nodes of parallel GA. In addition, the current MI-BCI pattern recognition was conducted in offline EEG signals. However, online MI-BCI still has many problems that affect feature precision and algorithm performance that need to be solved urgently. For the future works, the proposed algorithm will be applied to online MI-BCI for stroke treatment. Although the proposed algorithm can learn specific features in cross-subject and cross-session situations, it is obviously not advisable to train a recognition model for each subject. Future work also includes applying the proposed algorithm to transfer learning to build a cross-subject pattern recognition model.

7. Funding**^a

This work was supported by National Natural Science Foundation of China (Grant No.62106049); Natural Science Foundation of Fujian Province of China (Grant No 2022J01655).

8. Data availability**^a

The public benchmark motor imagery datasets can be downloaded at: <https://www.bbci.de/competition/iv/>.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

No potential conflict of interest was reported by the authors. We thank BCI competition IV for their share of motor imagery based EEG signals datasets. We thank the anonymous reviewers for their comments to improve the quality of the paper.

References

- [1] M.A. Lebedev, M.A. Nicolelis, Brain-machine interfaces: past, present and future, *Trends Neurosci.* 29 (9) (2006) 536–546.
- [2] M. Lebedev, Brain-machine interfaces: an overview, *Transl. Neurosci.* 5 (1) (2014) 99–110.
- [3] L. Tonin, J.D.R. Millán, Noninvasive brain-machine interfaces for robotic devices, *Ann. Rev. Control., Robot., Autonom. Syst.* 4 (1) (2021) 191–214.
- [4] A.B. Rapeaux, T.G. Constandinou, Implantable brain machine interfaces: first-in-human studies, technology challenges and trends, *Curr. Opin. Biotechnol.* 72 (2021) 102–111.
- [5] R.C. Mojoli, P.H.J. Nardelli, M.T. Barros, W. Saad, A. Hekmatmanesh, P.E.G. Silva, A.S. de Sena, M. Dzaferagic, H. Siljak, W. Van Leeuwijk, D.C. Melgarejo, S. Latre, Neurosciences and wireless networks: The potential of brain-type communications and their applications, *IEEE Commun. Surv. Tutorials* 23 (3) (2021) 1599–1621.
- [6] L. Tonin, E. Menegatti, D. Coyle, Advances in the Integration of Brain-Machine Interfaces and Robotic Devices, *Front. Robot. AI* 8 (2021) 22.
- [7] Z. Liu, J. Shore, M. Wang, F. Yuan, A. Buss, X. Zhao, A systematic review on hybrid EEG/fNIRS in brain-computer interface, *Biomed. Signal Process. Control* 68 (2021) 102595.
- [8] J. Mladenović, Standardization of protocol design for user training in EEG-based brain-computer interface, *J. Neural Eng.* 18 (1) (2021), 011003.
- [9] C. Ieracitano, N. Mammone, A. Hussain, F.C. Morabito, A novel explainable machine learning approach for EEG-based brain-computer interface systems, *Neural Comput. Appl.*, in press 34 (14) (2021) 11347–11360.
- [10] L. Xu, M. Xu, T.P. Jung, D. Ming, Review of brain encoding and decoding mechanisms for EEG-based brain-computer interface, *Cogn. Neurodyn.* 15 (4) (2021) 569–584.
- [11] D.B. Silversmith, R. Abiri, N.F. Hardy, N. Natraj, A. Tu-Chan, E.F. Chang, K. Ganguly, Plug-and-play control of a brain-computer interface through neural map stabilization, *Nat. Biotechnol.* 39 (3) (2021) 326–335.
- [12] D. Rathee, H. Raza, S. Roy, G. Prasad, A magnetoencephalography dataset for motor and cognitive imagery-based brain-computer interface, *Sci. Data* 8 (1) (2021) 1–10.
- [13] H. Gu, C.A. Chou, Detecting Epileptic Seizures via Non-Uniform Multivariate Embedding of EEG Signals, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE (2021, November) pp. 1690–1693.
- [14] A.J. Cook, K.J. Peifer, P.A. Tass, A Single Case Feasibility Study of Sensorimotor Rhythm Neurofeedback in Parkinson’s Disease, *Front. Neurosci.* 15 (2021), 623317.
- [15] D.N.T. Doan, B. Ku, J. Choi, M. Oh, K. Kim, W. Cha, J.U. Kim, Predicting dementia with prefrontal electroencephalography and event-related potential, *Front. Aging Neurosci.* 13 (2021) 180.
- [16] X.i. Zhao, Z. Wang, M. Zhang, H. Hu, A comfortable steady state visual evoked potential stimulation paradigm using peripheral vision, *J. Neural Eng.* 18 (5) (2021) 056021.
- [17] J. Decety, The neurophysiological basis of motor imagery, *Behav. Brain Res.* 77 (1–2) (1996) 45–52.
- [18] G. Pfurtscheller, EEG event-related desynchronization (ERD) and synchronization (ERS), *Electroencephalogr. Clin. Neurophysiol.* 1 (103) (1997) 26.
- [19] S. Chen, X. Shu, H. Wang, L. Ding, J. Fu, J. Jia, The Differences Between Motor Attempt and Motor Imagery in Brain-Computer Interface Accuracy and Event-Related Desynchronization of Patients With Hemiplegia, *Front. Neurorob.* 15 (2021), 706630.
- [20] S.B. Lee, H.J. Kim, H. Kim, J.H. Jeong, S.W. Lee, D.J. Kim, Comparative analysis of features extracted from EEG spatial, spectral and temporal domains for binary and multiclass motor imagery classification, *Inf. Sci.* 502 (2019) 190–200.
- [21] S. Lemm, B. Blankertz, G. Curio, K.R. Muller, Spatio-spectral filters for improving the classification of single trial EEG, *IEEE Trans. Biomed. Eng.* 52 (9) (2005) 1541–1548.
- [22] K.K. Ang, Z.Y. Chin, H. Zhang, C. Guan, Filter bank common spatial pattern (FBCSP) in brain-computer interface, in: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence) IEEE (2008, June), pp. 2390–2397.
- [23] S.H. Park, D. Lee, S.G. Lee, Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification, *IEEE Trans. Neural Syst. Rehabil. Eng.* 26 (2) (2017) 498–505.

- [24] H. Higashi, T. Tanaka, Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification, *IEEE Trans. Biomed. Eng.* 60 (4) (2012) 1100–1110.
- [25] Y. Zhang, G. Zhou, J. Jin, X. Wang, A. Cichocki, Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface, *J. Neurosci. Methods* 255 (2015) 85–91.
- [26] Y. Park, W. Chung, Frequency-optimized local region common spatial pattern approach for motor imagery classification, *IEEE Trans. Neural Syst. Rehabil. Eng.* 27 (7) (2019) 1378–1388.
- [27] A.S. Aghaei, M.S. Mahanta, K.N. Plataniotis, Separable common spatio-spectral patterns for motor imagery BCI systems, *IEEE Trans. Biomed. Eng.* 63 (1) (2015) 15–29.
- [28] J. Meng, G. Huang, D. Zhang, X. Zhu, Optimizing spatial spectral patterns jointly with channel configuration for brain-computer interface, *Neurocomputing* 104 (2013) 115–126.
- [29] V. Peterson, D. Wyser, O. Lambercy, R. Spies, R. Gassert, A penalized time-frequency band feature selection and classification procedure for improved motor intention decoding in multichannel EEG, *J. Neural Eng.* 16 (1) (2019) 016019.
- [30] P. Gaur, H. Gupta, A. Chowdhury, K. McCreadie, R.B. Pachori, H. Wang, A sliding window common spatial pattern for enhancing motor imagery classification in EEG-BCI, *IEEE Trans. Instrum. Meas.* 70 (2021) 1–9.
- [31] Y. Zhang, C.S. Nam, G. Zhou, J. Jin, X. Wang, A. Cichocki, Temporally constrained sparse group spatial patterns for motor imagery BCI, *IEEE Trans. Cybern.* 49 (9) (2018) 3322–3332.
- [32] V. Mishuhina, X. Jiang, Complex common spatial patterns on time-frequency decomposed EEG for brain-computer interface, *Pattern Recogn.* 115 (2021) 107918.
- [33] F. Lotte, C. Guan, Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms, *IEEE Trans. Biomed. Eng.* 58 (2) (2010) 355–362.
- [34] F. Yger, M. Berar, F. Lotte, Riemannian approaches in brain-computer interfaces: a review, *IEEE Trans. Neural Syst. Rehabil. Eng.* 25 (10) (2016) 1753–1762.
- [35] J.S. Kirar, R.K. Agrawal, A combination of spectral graph theory and quantum genetic algorithm to find relevant set of electrodes for motor imagery classification, *Appl. Soft Comput.* 97 (2020) 105519.
- [36] O.P. Idowu, O. Adelopo, A.E. Ilesanmi, X. Li, O.W. Samuel, P. Fang, G. Li, Neuro-evolutionary approach for optimal selection of EEG channels in motor imagery based BCI application, *Biomed. Signal Process. Control* 68 (2021) 102621.
- [37] M. Leon, C. Parkkila, J. Tidare, N. Xiong, E. Astrand, (June). Impact of NSGA-II objectives on EEG feature selection related to motor imagery, in: Proceedings of the 2020 Genetic and Evolutionary Computation Conference, 2020, pp. 1134–1142.
- [38] L. Zhang, D. Wen, C. Li, R. Zhu, Ensemble classifier based on optimized extreme learning machine for motor imagery classification, *J. Neural Eng.* 17 (2) (2020), 026004.
- [39] Y. Li, L. Wu, T. Wang, N. Gao, Q. Wang, EEG signal processing based on genetic algorithm for extracting mixed features, *Int. J. Pattern Recognit. Artif. Intell.* 33 (06) (2019) 1958008.
- [40] J. Tidare, M. Leon, E. Astrand, Time-resolved estimation of strength of motor imagery representation by multivariate EEG decoding, *J. Neural Eng.* 18 (1) (2021) 016026.
- [41] A. Vasilyev, S. Liburkina, L. Yakovlev, O. Perepelkina, A. Kaplan, Assessing motor imagery in brain-computer interface training: psychological and neurophysiological correlates, *Neuropsychologia* 97 (2017) 56–65.
- [42] K.K. Ang, Z.Y. Chin, C. Wang, C. Guan, H. Zhang, Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b, *Front. Neurosci.* 6 (2012) 39.
- [43] S. Aydin, A new combination: scale-space filtering of projected brain activities, *Med. Biol. Eng. Compu.* 47 (4) (2009) 435–440.
- [44] S. Aydin, S. Demirtaş, M.A. Tunga, K. Ateş, Comparison of hemispheric asymmetry measurements for emotional recordings from controls, *Neural Comput. Appl.* 30 (4) (2018) 1341–1351.
- [45] B. Kılıç, S. Aydin, Classification of Contrasting Discrete Emotional States Indicated by EEG Based Graph Theoretical Network Measures, *Neuroinformatics*, in press 20 (4) (2022) 863–877.
- [46] M.S. Treder, H. Purwins, D. Miklody, I. Sturm, B. Blankertz, Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification, *J. Neural Eng.* 11 (2) (2014) 026009.
- [47] J. Li, S.R. Wu, X. Zhang, T.J. Luo, R. Li, Y. Zhao, H. Peng, Cross-subject aesthetic preference recognition of Chinese dance posture using EEG, *Cogn. Neurodyn.*, in press (2022) 1–19.
- [48] R. Li, R. Qin, J. Zhang, J. Wu, C. Zhou, The esthetic preference of Chinese typefaces—An event-related potential study, *Brain Res.* 1598 (2015) 57–65.
- [49] H. Dong, T. Li, R. Ding, J. Sun, A novel hybrid genetic algorithm with granular information for feature selection and optimization, *Appl. Soft Comput.* 65 (2018) 33–46.
- [50] D. Rajeswari, M. Prakash, J. Suresh, Computational grid scheduling architecture using MapReduce model-based non-dominated sorting genetic algorithm, *Soft. Comput.* 23 (18) (2019) 8335–8347.
- [51] J. Wang, Z. Feng, X. Ren, N.a. Lu, J. Luo, L. Sun, Feature subset and time segment selection for the classification of EEG data based motor imagery, *Biomed. Signal Process. Control* 61 (2020) 102026.
- [52] P.S. Bradley, O.L. Mangasarian, Feature selection via concave minimization and support vector machines, in: ICML, Vol. 98, (1998, July), pp. 82–90.
- [53] D. Wang, D. Miao, G. Blohm, Multi-class motor imagery EEG decoding for brain-computer interfaces, *Front. Neurosci.* 6 (2012) 151.
- [54] I. Güler, E.D. Übeyli, Multiclass support vector machines for EEG-signals classification, *IEEE Trans. Inf Technol. Biomed.* 11 (2) (2007) 117–126.
- [55] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, *ACM Trans. Intell. Syst. Technol. (TIST)* 2 (3) (2011) 1–27.
- [56] M. Tangermann, K.R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, ... & B. Blankertz, Review of the BCI competition IV, *Front. Neurosci.* 55 (2012).
- [57] V. Mishuhina, X. Jiang, Feature weighting and regularization of common spatial patterns in EEG-based motor imagery BCI, *IEEE Signal Process Lett.* 25 (6) (2018) 783–787.
- [58] H. Bashashati, R.K. Ward, A. Bashashati, User-customized brain computer interfaces using Bayesian optimization, *J. Neural Eng.* 13 (2) (2016) 026001.
- [59] X. Jia, Y. Song, L. Yang, L. Xie, Joint spatial and temporal features extraction for multi-classification of motor imagery EEG, *Biomed. Signal Process. Control* 71 (2022) 103247.
- [60] R.T. Schirrmeister, J.T. Springenberg, L.D.J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, T. Ball, Deep learning with convolutional neural networks for EEG decoding and visualization, *Hum. Brain Mapp.* 38 (11) (2017) 5391–5420.
- [61] V.J. Lawhern, A.J. Solon, N.R. Waytowich, S.M. Gordon, C.P. Hung, B.J. Lance, EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces, *J. Neural Eng.* 15 (5) (2018) 056013.
- [62] S.U. Amin, M. Alsulaiman, G. Muhammad, M.A. Mekhtiche, M.S. Hossain, Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion, *Future Gener. Comput. Syst.* 101 (2019) 542–554.
- [63] S. Sakhavi, C. Guan, S. Yan, Learning temporal information for brain-computer interface using convolutional neural networks, *IEEE Trans. Neural Networks Learn. Syst.* 29 (11) (2018) 5619–5629.
- [64] E. Huang, X. Zheng, Y. Fang, Z. Zhang, Classification of Motor Imagery EEG Based on Time-Domain and Frequency-Domain Dual-Stream Convolutional Neural Network, *IRBM* 43 (2) (2021) 107–113.
- [65] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, *J. Mach. Learn. Res.* 9 (11) (2008) 1–12.