



A new parameter tuning approach for enhanced motor imagery EEG signal classification

Shiu Kumar^{1,2} · Alok Sharma^{2,3,4} 

Received: 22 October 2017 / Accepted: 19 March 2018
© International Federation for Medical and Biological Engineering 2018

Abstract

A brain-computer interface (BCI) system allows direct communication between the brain and the external world. Common spatial pattern (CSP) has been used effectively for feature extraction of data used in BCI systems. However, many studies show that the performance of a BCI system using CSP largely depends on the filter parameters. The filter parameters that yield most discriminating information vary from subject to subject and manually tuning of the filter parameters is a difficult and time-consuming exercise. In this paper, we propose a new automated filter tuning approach for motor imagery electroencephalography (EEG) signal classification, which automatically and flexibly finds the filter parameters for optimal performance. We have evaluated the performance of our proposed method on two public benchmark datasets. Compared to the existing conventional CSP approach, our method reduces the average classification error rate by 2.89% and 3.61% for BCI Competition III dataset IVa and BCI Competition IV dataset I, respectively. Moreover, our proposed approach also achieved lowest average classification error rate compared to state-of-the-art methods studied in this paper. Thus, our proposed method can be potentially used for developing improved BCI systems, which can assist people with disabilities to recover their environmental control. It can also be used for enhanced disease recognition such as epileptic seizure detection using EEG signals.

Keywords Brain-computer interface (BCI) · Filter tuning · Genetic algorithm (GA) · Motor imagery (MI) · Temporal filters

1 Introduction

A brain-computer interface (BCI) system can bridge the gap between humans and computers by translating thoughts into control signals, which can be used to control external devices

[24]. It can be used to assist severely disabled people to live a quality life with minimum or no dependence. Motor imagery (MI) is an important paradigm in developing a BCI system that allows direct human-computer interaction (HCI), i.e., allowing communication between the brain and the external devices without the involvement of peripheral nerves and muscles [24]. BCI research is increasingly gaining interest as it aims to re-establish independence and reduce social exclusion for people with disabilities. BCI systems have been explored for a wide range of applications such as communication control [12], movement control [20], environment control [5], neuro-rehabilitation [9, 22, 30, 31, 43], P300-based word typing system [1], biometric identification [18], and brain wave-controlled robots and wheelchairs for the disabled people [23, 28]. Biomedical engineering is the major focus and emphasis of the current BCI research [9, 14–16, 21, 22, 30, 42, 44, 51, 53, 58].

Research in the field of MI-BCI systems is ongoing as there are still many problems that exist and are yet to be addressed properly. Focus is also on developing more accurate, reliable, and computationally efficient MI-BCI systems that can be incorporated with wearable devices. A range of approaches have been reported for solving different aspects of the problem.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11517-018-1821-4>) contains supplementary material, which is available to authorized users.

✉ Alok Sharma
alok.sharma@griffith.edu.au

¹ Department of Electronics, Instrumentation & Control Engineering, School of Electrical & Electronics Engineering, Fiji National University, Samabula, Fiji

² School of Engineering and Physics, Faculty of Science, Technology & Environment, The University of the South Pacific, Suva, Fiji

³ Institute for Integrated and Intelligent Systems (IIIS), Griffith University, Brisbane, Australia

⁴ Laboratory for Medical Science Mathematics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan

CSP has been widely used in MI-BCI applications as it aims to find the spatial filters that maximize the variance of one class, while minimizing the variance of the other class in order to discriminate between different classes. CSP was first introduced for a two-class problem and later extended to multi-class problem. The conventional CSP approach uses a single wide frequency band mostly in the range of 8–30 Hz, covering both the mu and beta rhythms. However, certain frequency components are useless resulting in degrading the overall performance of the system. To improve the performance of the conventional CSP approach, common spatio-spectral pattern (CSSP) [19] and common sparse spectral spatial pattern (CSSSP) [7] methods were proposed.

The responsive frequency bands vary from subject to subject, and to address this problem, the use of multiple frequency bands has been proposed [3, 25, 46]. In [25], a sub-band CSP (SBCSP) method has been proposed. This paper uses multiple filters (filter banks) to decompose the signal into multiple bands called sub-bands. CSP is then performed separately on each of these sub-bands to extract the variance-based CSP features. Linear discriminant analysis (LDA) is then performed on each feature set obtained from the separate sub-bands in order to reduce the feature dimensionality of each sub-band to one. Thus, one-dimensional features of each sub-band (LDA scores) are fused together and fed into a support vector machine (SVM) classifier. The method performed well and outperformed the CSP, CSSP, and CSSSP methods. The shortfall of the SBCSP method is that since a number of sub-bands are used, the number of extracted features also increased and some of which are not useful and tend to degrade the performance of the system. To address this issue, filter bank CSP (FBCSP) [3] has been proposed, in which all the variance-based CSP features from each sub-band are combined together and different feature selection and classification methods have been evaluated. Feature selection is employed in order to select the significant features, thereby removing the redundant features. Although FBCSP outperformed the SBCSP method, it still utilized several sub-bands that lead to higher computational cost. To address this issue, a discriminative filter bank CSP (DFBCSP) [46] method has been proposed. In DFBCSP, the signals are decomposed into multiple sub-bands (12 sub-bands in the frequency range of 6–40 Hz) during training phase. Fisher's ratio of the band spectral power of channel C3 is then used to select the top 4 sub-bands that give maximum discrimination between the different tasks. Channel C3 located on the contralateral hemisphere was selected as signals from sensorimotor cortex have shown to have maximum discriminating power amongst several MI tasks. Signals from contralateral channels C3, C4, and Cz and its surrounding channels were used for estimating Fisher's ratio. The authors discovered that single channel C3 or C4 alone offered superior performance in selecting the sub-bands and thus channel C3 has been used. The selected sub-bands are then used for extracting the variance-based CSP features and fed into

an SVM classifier for classification. The method achieved a reduced computational cost compared to SBCSP and FBCSP and also performed better.

To address the issue of computation cost due to the use of large number of sub-bands and due to the fact that some sub-bands may contain redundant information, a sub-band binary particle swarm optimization (BPSO) approach using CSP [49] has been proposed. In sub-band BPSO-CSP, the EEG signals are decomposed into sub-bands using multiple filter banks as in DFBCSP. Then, the BPSO algorithm is employed to select the sub-bands that give optimal performance using the training data. Recently, a sparse filter band CSP (SFBCSP) [56] method that uses multiple filter bands is proposed, which optimizes the sparse patterns. Bayesian learning has also gained increased attention recently and has been used for feature selection in various applications [45, 57]. In [54], a sparse Bayesian learning approach (SBLFB) has been used for obtaining sparse features that are used for MI EEG signal classification. A number of studies have explored feature optimization and classifier optimization for improved EEG signal classification in BCI applications. A method based on multi-kernel extreme learning machine (MKELM) [55] has been proposed for EEG signal classification. Gaussian and polynomial kernels have been integrated together with multi-kernel learning approach for more robust classification of multiple non-linear feature spaces. In [48], a spatio-temporal feature extraction in conjunction with multivariate linear regression has been proposed, which improved the classification performance by learning discriminative steady-state visual evoked potential (SSVEP) features. A number of component analysis methods have also been proposed for processing biomedical data such as independent component analysis (ICA), blind source separation, and matrix-based component analysis methods [58].

During MI, the power of the EEG signal in certain specific regions of the brain decreases and later increases as the MI concludes. These physiological phenomena are known as event-related desynchronization (ERD) and event-related synchronization (ERS), respectively [26]. ERD and ERS are directly related to sensory motor rhythms mu and beta. The mu (8–12 Hz) and beta (18–25 Hz) rhythms are excellent features for MI signals. However, the frequency band varies across different subjects together with the varying psychological states of the subjects [2]. CSP has been widely explored and used effectively for formulating spatial filters in ERD/ERS detection [29]. Although CSP has performed considerably well for spatial feature extraction, the filtering of signals in the time domain confines its performance. Using improper frequency range for the filter or simply using unfiltered EEG signal generally yields undesirable classification results [25]. Thus, while employing CSP, usually a broad frequency band is used or the filters are manually tuned for each of the subjects [17]. These approaches are not suitable for obtaining optimal performance as the wide band contains certain frequency

ranges that are not useful. On the other hand, manually tuning is a time-consuming and difficult approach. When filtering is performed on raw data, the filter parameters need to be selected carefully [50].

To address the problem of selecting subject-specific frequency band that achieves optimal performance using the CSP algorithm, this paper proposes the temporal filter parameter optimization with CSP approach (TFPO-CSP predictor). A BCI system usually involves four phases: (1) temporal filtering of the EEG data, (2) feature extraction, (3) feature selection and (4) classification. In this work, we mainly focused on phase 1, i.e., temporal filtering to obtain the optimum parameters and the other 3 phases will be implemented as in CSP approach. We also aimed to employ a single bandpass filter in a quest to keep the computation complexity of the resulting system to a minimum. Generally, in phase 1, the signal is filtered either using single or multiple bandpass filters. Spatial filtering is performed and features are extracted in phase 2. The third phase is usually incorporated to select the most discriminating features in order to improve the system performance. The final phase involves using a suitable classifier to classify the EEG signals. The public benchmark dataset IVa of BCI Competition III and dataset I of BCI Competition IV are used to show that the proposed approach can be used to select the appropriate subject-specific frequency band for optimal performance. The effectiveness of the proposed TFPO-CSP predictor over the state-of-the-art methods such as FBCSP, DFBCSP, and SBLFB is demonstrated.

The following sections of the paper are organized as follows. Section 2 presents the proposed TFPO-CSP predictor while in Sections 3 and 4 the experimental study and findings are presented and discussed. Section 5 draws the conclusion and gives insight of some future works.

2 Methods

2.1 Description of the datasets

The public benchmark Dataset IVa of BCI Competition III [6] and Dataset I of BCI Competition IV [4] (referred to as dataset 1 and dataset 2 from here onwards, respectively) have been used for evaluating the effectiveness of the proposed approach.

Dataset 1 contains 118 channels of EEG signals for right hand and left foot MI tasks, which have been recorded from five healthy subjects labeled *aa*, *al*, *av*, *aw*, and *ay*. The down sampled signal at 100 Hz has been used. It contains 140 trials of each task for each of the subjects. A detail description of the dataset can be found online at [<http://www.bbc.de/competition/iii/>].

Dataset 2 contains 59 channels of real long-term EEG signals recorded for left hand and right hand MI tasks acquired from 7 healthy subjects (named *a* to *g*). The down sampled signal at 100 Hz has been used, which contains 200 trials for each subject with equal number of each type of MI tasks. A

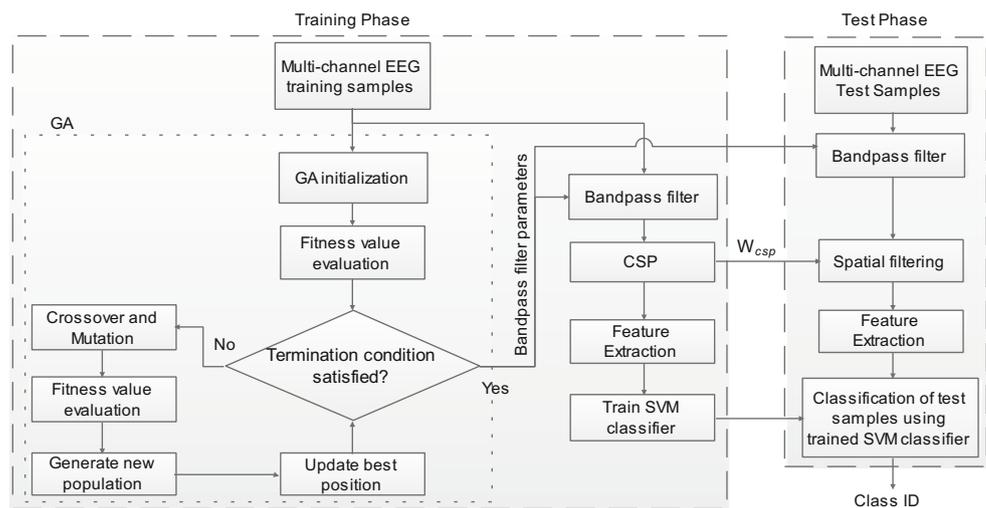
detail description of the dataset can be found online at [<http://www.bbc.de/competition/iv/>].

2.2 Experimental setup

In this study, we have extracted the EEG data between 0.5 and 2.5 s (i.e., 200 time points) after the visual cue. The number of CSP filters is set to 6 (i.e., $m = 3$) in all experiments, except for the DFBCSP and SBLFB methods in which two CSP filters have been used (i.e., $m = 1$). In all experiments (except for the proposed approach), an eighth-order butterworth bandpass filter has been used and SVM is used as the classifier. The following experimental settings have been used for each of the methods:

- *CSP*: A butterworth bandpass filter with passband of 7–30 Hz has been applied.
- *CSSP*: Bandpass filter is the same as in CSP. The sample point delay τ value was chosen from 1 to 15 (as in [8]), which was selected by performing 10-fold cross-validation on the train data. The average error rate of the 10-fold cross-validation for each value of τ is calculated and the value of τ giving the minimum error rate is selected for testing.
- *FBCSP*: 6 bandpass filters having bandwidth of 6 Hz in the range of 4 to 40 Hz with no overlap have been used. These were adopted from Higashi and Tanaka [8], as using these settings gave optimal results. Mutual information-based feature selection has been performed as it gave the best results in [3].
- *DFBCSP*: 12 bandpass filters with bandwidth of 4 Hz in the range of 6 to 40 Hz have been used, as described in [46]. The top four bands were selected using Fisher's ratio and further utilized for classification.
- *SFBCSP*: 17 bandpass filters having bandwidth of 4 Hz with an overlap of 2 Hz in the range of 4–40 Hz have been used, as in [56]. The regularization parameter λ was chosen using 10-fold cross-validation.
- *SBLFB*: As in [54], 17 bandpass filters having bandwidth of 4 Hz with an overlap of 2 Hz in the range of 4–40 Hz have been used. Multiple sub-band features are learned by SBLFB.
- *BPSO-CSP*: 10 bandpass filters having bandwidth of 4 Hz in the range of 8 to 30 Hz with an overlap of 2 Hz have been used. Only 25 selected channels of data were used for processing as in Wei and Wei [49].
- *TFPO-CSP (PSO)*: This method uses particle swarm optimization (PSO) as the optimization algorithm. The number of particles for the PSO algorithm is set to 10 because with 10 particles, the optimal solution is found in a reasonable amount of time. The dimension of each particle is set to 3, where dimensions represent the lower cutoff frequency, upper cutoff frequency, and filter order of the butterworth bandpass filter. The range of the dimensions was set as follows: lower cutoff frequency [0.50, 16.0 Hz], upper cutoff frequency [18, 32 Hz], and filter order [1, 30]. The frequency ranges were selected so

Fig. 1 Framework for the proposed TFPO-CSP predictor with GA as the optimization algorithm



as to include the frequencies of the mu and beta rhythms. The range for the velocity was set to $[-4, 4]$. The maximum number of iterations (I_{max}) was set to 35, which was decided after carrying out a number of experiments. In TFPO-CSP (PSO), different subjects have different fitness condition values. Thus, several trial runs were conducted without the use of fitness condition in order to determine the minimum error rate that can be achieved by the TFPO-CSP (PSO) for each subject. It was noted that the minimum error rates that could be obtained for each subject were similar to the best error rate that can be achieved amongst the DFBCSP and SBLFB methods. Therefore, the fitness condition value used for each subject was set to the minimum of the error rates obtained amongst the DFBCSP and SBLFB methods.

- **TFPO-CSP (GA):** This method uses genetic algorithm (GA) as the optimization algorithm. The frequency ranges, filter order, maximum number of iterations, and fitness conditions were same as those used for TFPO-CSP (PSO). The population size was set to 10, which is same as the number of particles used for TFPO-CSP (PSO). Tournament selection has been employed with tournament size of 3. A crossover percentage of 0.7 and mutation percentage of 0.3 have been used. Fitness-based survivor selection is performed. For BCI Competition III dataset IVa, the fitness conditions were set to 10, 1.5, 25, 5.0, and 4.5 for subjects *aa*, *al*, *av*, *aw*, and *ay*, respectively, while for BCI Competition IV dataset I, the fitness conditions were set to 13, 41.5, 33, 21, 11, 14, and 8 for subjects *a* to *g*, respectively.
- **TFPO-CSP (GA)*:** All experimental settings were same as TFPO-CSP (GA), except that only 25 selected channels of data as used in BPSO-CSP were used for processing in order to compare our proposed method with BPSO-CSP.
- **TFPO-CSP (ABC):** This method uses artificial bee colony (ABC) as the optimization algorithm. The frequency ranges, filter order, maximum number of iterations, and fitness conditions were same as those used for TFPO-CSP (PSO). The colony size was set to 10, which is same as the number of particles used for TFPO-CSP (PSO).

2.3 Proposed TFPO-CSP method

The framework for the proposed TFPO-CSP predictor is presented in Fig. 1. The temporal filtering of the MI EEG signal is integral because CSP mostly depends on responsive frequency band in order to perform well. Unfiltered signal or filtering the signal with inappropriate frequency band would result in poor performance of the system. Moreover, the responsive frequency band varies from subject to subject and thus using a fixed frequency band is detrimental in achieving a system with optimal performance. Therefore, in order to find the optimum subject-specific responsive frequency band, GA has been employed for finding optimal filter band parameters. The proposed TFPO-CSP predictor involves four phases: temporal filter parameter selection and filtering, spatial filtering, feature extraction, and classification.

The first phase that involves selecting the temporal filter parameters using GA is the major focus of this paper as it aims to find a single responsive filter, which can give optimal results. The GA has been modified and used with our proposed system in order to select the best parameters (upper cutoff frequency, lower cutoff frequency, and filter order) of a bandpass butterworth filter. Once the filter parameters are selected using GA, the raw EEG signal is bandpass filtered using the selected filtering parameters. In the second phase, CSP is utilized for spatial filtering. The variance-based CSP features are then extracted from the spatially filtered data during the third phase, while classification using SVM classifier is performed in the final phase. The following sub-sections present each of the phases in more detail.

The ABC [10] and PSO [27] algorithms have also been modified to be used as the optimization algorithm in our proposed predictor for finding the optimal parameters. Details of the ABC and PSO algorithms are given in [supplementary materials](#).

2.3.1 Optimization of temporal filter parameters

For optimizing the parameters of the temporal filter, we have employed GA [52]. GA is based on natural selection and

genetics and is mostly used for optimization and search problems. A population of chromosomes (solutions) is generated with dimension equal to the number of parameters to be optimized. This can be done either randomly or by heuristic initialization. The first generation is referred to as the parents. Children are then generated from selected parents using cross-over. The selection of parents can be done using various methods such as Roulette wheel selection, stochastic universal sampling (SUS), tournament selection, rank selection, and random selection. Mutation is then performed on a number of randomly selected parents to form new mutants. Finally, survivor selection is done and the chromosomes that survive become the parents for the next iteration. This process is repeated until the desired fitness condition or the maximum number of iterations is reached. The pseudo code of the proposed framework is given in Algorithm 1 and 2. The fitness function value referred to in Algorithm 2 is the error returned by Algorithm 1.

Algorithm 1 Pseudo code for fitness function algorithm

Function fitnessFunction (*GBest*, *trainSamples*, *trainLabels*, *index*)

Input: *GBest* - global best position of all particles
trainSamples - N_p labeled training samples
trainLabels - N_c train sample classes
index - cross-validation indices

Output: average of all the errors

initialization:

1. Initialize f_s (sampling frequency)
2. Bandpass filter *trainSamples* using parameters from *GBest*
3. $itr = 0$

while $itr \leq \text{number of folds for cross-validation}$

1. *test* = indices in *index* equal to *itr*
2. *train* = indices in *index* apart from those in *test*
3. Determine W_{CSP} using *trainSamples* at *train* indices
4. Spatially filter *trainSamples* using (1)
5. Extract features using (2)
6. Train classifier using features extracted from *trainSamples* at *train* indices
7. $error(itr)$ = classification error using features extracted from *trainSamples* at *test* indices (using trained classifier)
8. $itr++$

end while

return $\text{average}(error)$

2.4 CSP feature extraction

The CSP technique has gained a lot of attention and has been widely used in recent years for MI-BCI systems since its inception, where it has been used for detection of abnormalities in EEG signal [13]. CSP projects the data to a new time series maximizing the variance of one class, while minimizing the variance of the other class. A detailed explanation of the CSP algorithm can be found in our previous work [17]. Consider the bandpass-filtered EEG sample $X_n \in R^{C \times T}$, where n denotes the n th sample, C is the number of channels and T is the number of time points. The spatially filtered signal Z_n can be obtained using

(1), where W_{CSP} is the CSP spatial filter that is formed by selecting the first and last m columns of CSP matrix, W .

Algorithm 2 Pseudo code for GA algorithm

Function GA($N_p, I_{max}, ur, lr, FC, trainSamples, trainLabels$)

Inputs: N_p - population size;
 I_{max} - maximum number of iterations
 ur - the upper range of the parameters
 lr - the lower range of the parameters
 FC - fitness condition
trainSamples - N_p training samples
trainLabels - N_c train sample classes

Output: *GBest* (global best position)

Initialization (GA parameters):

1. Initialize *cp* (crossover percentage)
2. Initialize *gamma*
3. Initialize *mp* (mutation percentage)
4. Initialize *tr* (tournament size)
5. Initialize number of offsprings (parents)
 $nc = 2 * \text{round}(cp * N_p / 2)$
6. Initialize number of mutants, $nm = \text{round}(mp * N_p)$
7. Randomly initialize the positions of the population (*chromosomes*, C) within the upper and lower range of the parameters
8. Evaluate **fitness function** and initialize fitness value to be the best position of $C(i)$
9. Initialize *GBest*, *GBest_value* (best fitness value), *GWorst* (chromosome with worst fitness value)

While $itr < I_{max}$ **and** $GBest_value > FC$

for $i = 1:nc/2$ (half of number of parents)

1. Use tournament selection to choose 2 parents ($p1$ & $p2$)
2. Perform crossover using whole arithmetic recombination:
 $c_1 = \alpha * p_1 + p_2(1 - \alpha)$
 $c_2 = \alpha * p_2 + p_1(1 - \alpha)$
 α is a random number between $[-\text{gamma}, 1 + \text{gamma}]$
3. If the parameters of the children are out of range, limit them to the upper and lower range
4. Evaluate **fitness function** of c_1 and c_2

end for

for $i = 1:nm$

5. Randomly select a chromosome, C_i from the population
6. Let the mutant, $M_i = C_i$ and perform mutation as follows:
 $m_i = * C_{i,j} + \text{sigma}_j * r$, where r is a random number between $[0, 1]$ and $\text{sigma} = 0.1 * (ur_j - lr_j)$
7. If the parameters of the mutant are out of range, limit them to the upper and lower range
8. Evaluate **fitness function** of M_i

end for

9. Select N_p best chromosomes from the population, generated children's and mutants to be the new population
10. Set *GBest* to the position of the best chromosome
11. $itr++$

end while

return *GBest*

$$Z_n = W_{CSP}^T X_n \tag{1}$$

The variance-based CSP features of n th sample is then extracted using (2), where F_n^i is the i th feature of the n th sample, and $\text{var}(Z_n^j)$ denotes the variance of j th row of Z_n .

$$F_n^i = \log \left(\frac{\text{var}(Z_n^i)}{\sum_{j=1}^{2m} \text{var}(Z_n^j)} \right) \tag{2}$$

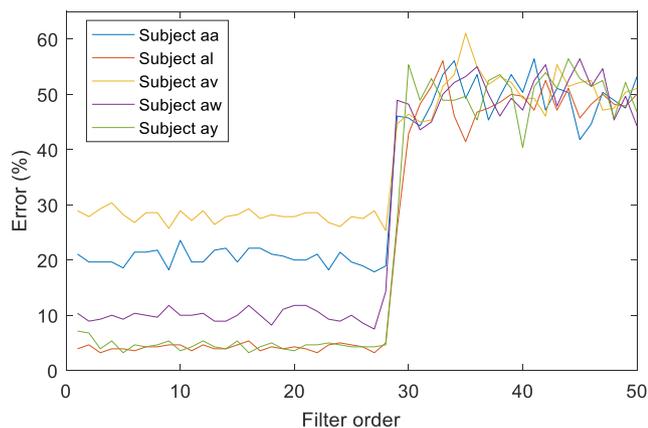


Fig. 2 Average error rates of 10-fold cross-validation with different values of filter order (results obtained using dataset 1)

Using the extracted features of each sample, the feature set is constructed using (3), where N is the number of samples. Thus, the features of training and test samples are used to construct the training and test feature sets, respectively.

$$F = \begin{bmatrix} F_1^1 & F_1^2 & \dots & F_1^{2m} \\ \vdots & \vdots & \ddots & \vdots \\ F_N^1 & F_N^2 & \dots & F_N^{2m} \end{bmatrix} \quad (3)$$

2.5 Classification

An SVM classifier with linear kernel is trained using the training feature set. The test feature set is then used to classify each of the test samples using the trained SVM classifier model.

Data availability For the benefit of the research community, the TFPO-CSP algorithm (MATLAB m-file) is available on our website: <http://www.alok-ai-lab.com/publications.php>.

3 Results

3.1 Evaluation scheme

The 10×10 -fold cross-validation scheme is used to evaluate the performance of all experiments conducted. The trials are divided into 10 sets. One set is selected as test set while all others are selected as training sets. Only the training set is used by GA algorithm to find the best temporal filter parameters. The fitness function again divides the training set of GA algorithm into 10 sets for performing 10-fold cross-validation. The average of these cross-validation values is returned by the fitness function, i.e., the average classification error rate. Thus, the GBest chromosome returned by GA is used to test the performance of the initial test set. This procedure is repeated 10 times with a different set used as test data and remaining used as training data of GA. The whole procedure was repeated 10 times and the average of these classification error rates is presented as results.

3.2 Experimental result analysis

The proposed TFPO-CSP predictor searches for the optimal frequency band and filter order. It can be observed (and suggested in the literature) that the responsive frequency bands vary between subjects, and thus different band leads to different outcome. It is therefore important to give due consideration to band optimization. Thus, in order to improve the performance, one aspect is to evaluate bands using the training dataset. We have evaluated the effect of filter order on the performance of the BCI system in Fig. 2. It shows the 10-fold cross-validation error rate for all the subjects of dataset 1. It can be observed from Fig. 2 that the filter order has a considerable effect on the overall error rate. Similar results were obtained for dataset 2 (refer to Fig. S1 of

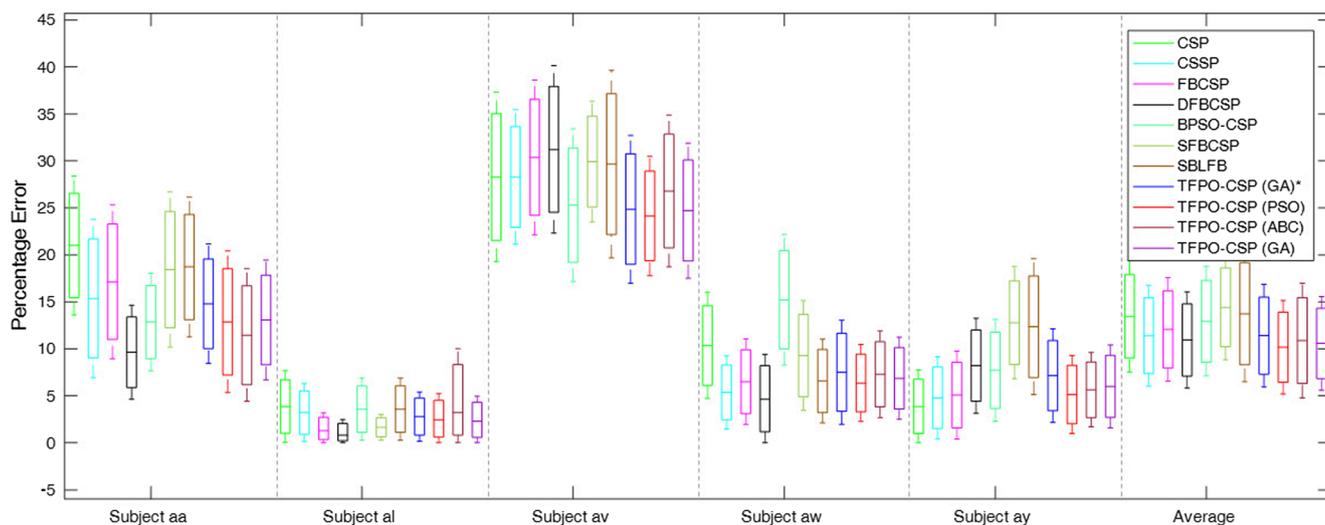


Fig. 3. 10×10 -fold cross-validation error rates for different methods using public benchmark dataset 1

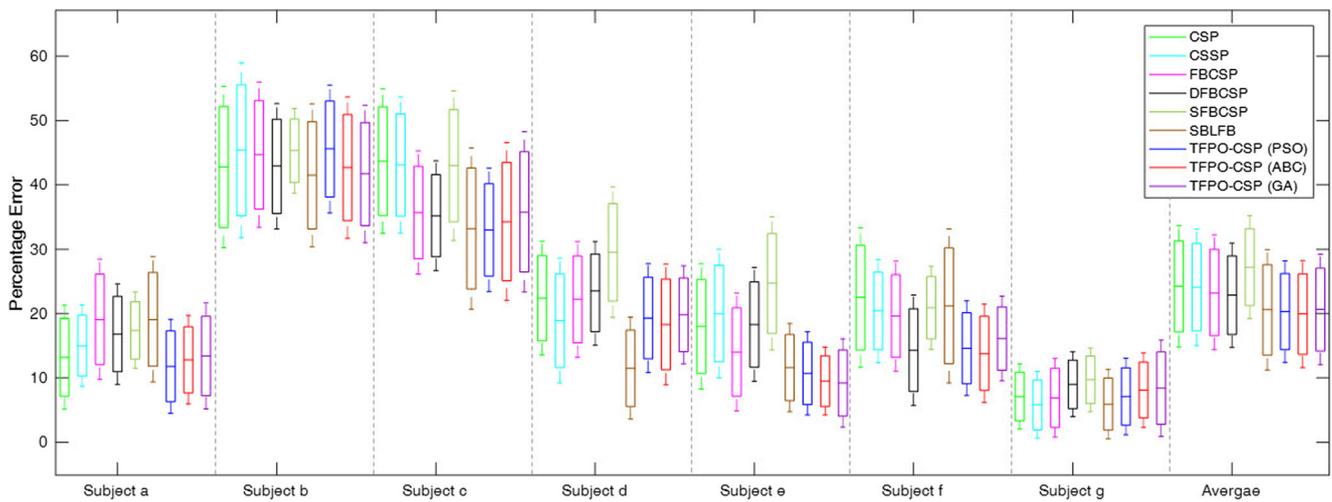


Fig. 4. 10 × 10-fold cross-validation error rates for different methods using public benchmark dataset 2

supplementary material). This explains the reason for optimizing the filter order together with the frequency band. Since the classification error rates are high for larger values of filter order, we have limited the filter order in the range between 1 and 30.

Two separate experiments were carried out to evaluate the performance of the proposed system. In the first experiment, we selected a subset of 25 channels (as in [49]) that are instrumental for the neurophysiological discrimination between the tasks. In the second experiment, all channels of data were used. The results obtained using dataset 1 and dataset 2 are shown in Figs. 3 and 4, respectively. The TFPO-CSP (GA)* indicates the results obtained when selected subset of 25 channels is used, while the TFPO-CSP (GA) indicates the result when all channels data are used for processing.

The results obtained show that the proposed TFPO-CSP predictor outperformed state-of-the-art methods studied in this paper in terms of the average classification error rates using either of the three optimization algorithms (PSO, GA, or ABC). The TFPO-CSP (ABC) and TFPO-CSP (PSO)

methods obtained average error rates that were within ± 1% of the average error rate obtained using TFPO-CSP (GA). Furthermore, our proposed method also performed better than the MKELM approach that achieved average classification error rate of 12.50%, evaluated on dataset 1. A spatial-frequency-temporal optimized feature sparse representation-based classification (SFTOFSRC) method is proposed in [21] which achieved error rate of 23.89% on dataset 1. The SFTOFSRC method used multiple frequency bands and multiple time segments. It also employed channel selection. While the authors showed that their method improves performance when sparse regression classification (SRC) with dictionary optimization is used, it obtained a considerable higher classification error rate. This could be due to the pre-processing that may have been used. On the other hand, the TFPO-CSP method achieved average classification error rate of 10.19% with common averaging as the only pre-processing step.

Using only selected 25 channels, we were able to achieve promising results and obtained the minimum average

Table 1 Ranges of parameters that were selected by TFPO-CSP (PSO), TFPO-CSP (GA), and TFPO-CSP (ABC) approaches for the different subjects (dataset 1)

Parameter	Method	Subject				
		<i>aa</i>	<i>al</i>	<i>av</i>	<i>aw</i>	<i>ay</i>
Lower cutoff frequency (Hz)	TFPO-CSP (PSO)	11–12	4–11	6–9	8–11	4–10
	TFPO-CSP (ABC)	10–13	3–12	6–15	8–12	4–11
	TFPO-CSP (GA)	10–12	6–12	6–15	8–11	5–10
Upper cutoff frequency (Hz)	TFPO-CSP (PSO)	26–31	20–28	23–30	24–32	22–32
	TFPO-CSP (ABC)	23–32	19–32	24–32	21–31	20–32
	TFPO-CSP (GA)	27–32	20–32	25–31	19–31	22–32
Filter order	TFPO-CSP (PSO)	17–30	2–15	2–29	4–25	4–10
	TFPO-CSP (ABC)	1–21	1–30	1–26	3–22	1–28
	TFPO-CSP (GA)	9–28	1–29	1–26	2–21	1–20

Table 2 Ranges of parameters that were selected by TFPO-CSP (PSO), TFPO-CSP (GA), and TFPO-CSP (ABC) approaches for the different subjects (dataset 2)

Parameter	Method	Subject						
		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>
Lower cutoff frequency (Hz)	TFPO-CSP (PSO)	6–12	7–14	6–11	6–12	9–12	7–10	5–14
	TFPO-CSP (ABC)	8–12	6–14	5–11	6–14	8–11	6–10	2–15
	TFPO-CSP (GA)	7–11	7–13	6–10	8–13	9–12	6–10	4–15
Upper cutoff frequency (Hz)	TFPO-CSP (PSO)	21–32	22–31	20–31	24–30	24–32	22–30	23–31
	TFPO-CSP (ABC)	20–31	22–32	20–32	20–30	22–32	22–31	23–32
	TFPO-CSP (GA)	20–32	20–32	21–32	23–32	22–32	22–31	21–32
Filter order	TFPO-CSP (PSO)	2–25	4–28	4–22	1–16	3–19	6–24	1–21
	TFPO-CSP (ABC)	1–29	2–27	1–24	1–21	1–14	2–29	1–24
	TFPO-CSP (GA)	2–24	1–30	2–26	1–22	1–19	2–23	2–25

classification error rate in comparison with CSP, CSSP, FBCSP, and BPSO-CSP. It can be noted that our proposed predictor with 25 channels (TFPO-CSP (GA)*) also shows very promising results. Therefore, the proposed approach is also recommended over other competing methods when computation complexity is a requirement, where data from fewer channels could be used. For dataset 2, subjects *a* and *c* achieved the minimum error rate using TFPO-CSP (GA) while subjects *e* and *f* achieved the minimum error rates using TFPO-CSP (ABC). The minimum average error rate was obtained using TFPO-CSP (PSO) and TFPO-CSP (ABC) for dataset 1 and dataset 2, respectively. However, there was no significant difference compared to the average error rate of TFPO-CSP (GA) (less than 0.5% difference in error rate). The SBLFB method also performed well on dataset 2 with subjects *b*, *d*, and *g* achieving the minimum error rate. The average error rate is also comparable to that of the proposed method; however, it did not perform well on dataset 1.

Tables 1 and 2 show the ranges in which the lower and upper cutoff frequencies and filter orders of the butterworth bandpass filter were selected using the TFPO-CSP (PSO), TFPO-CSP (GA), and TFPO-CSP (ABC) approaches for dataset 1 and dataset 2, respectively. For dataset 1 and dataset 2, it can be noted that the lower cutoff frequency range for all the subjects is in the range of 4–15 Hz, the upper cutoff frequency is in the range of 19–32 Hz, and the filter order is in the range of 1–30. The range of values (for lower and upper cutoff frequencies and filter order) that were obtained using the TFPO-CSP (PSO), TFPO-CSP (ABC), and TFPO-CSP (GA) were similar. All the frequency ranges are close to the frequency range of 8–30 Hz that includes both the mu and beta rhythms. It is clearly shown that the responsive frequency bands vary from subject to subject as the range of frequency bands selected for different subjects varies.

The average, minimum, and maximum numbers of iterations required by the TFPO-CSP (PSO), TFPO-CSP (ABC), and TFPO-CSP (GA) approaches to select the temporal filter

Table 3 The average, minimum, and maximum numbers of iterations taken by TFPO-CSP (PSO), TFPO-CSP (GA), and TFPO-CSP (ABC) approaches to select the temporal filter parameters (dataset 1)

Iterations	Method	Subject				
		<i>aa</i>	<i>al</i>	<i>av</i>	<i>aw</i>	<i>ay</i>
Minimum	TFPO-CSP (PSO)	1	1	1	2	1
	TFPO-CSP (ABC)	3	1	1	1	1
	TFPO-CSP (GA)	1	1	1	1	1
Average	TFPO-CSP (PSO)	6	2	4	25	2
	TFPO-CSP (ABC)	9	2	8	8	2
	TFPO-CSP (GA)	10	1	2	3	2
Maximum	TFPO-CSP (PSO)	35	17	15	35	17
	TFPO-CSP (ABC)	14	9	14	35	6
	TFPO-CSP (GA)	35	1	8	15	12

Table 4 The average, minimum, and maximum numbers of iterations taken by TFPO-CSP (PSO), TFPO-CSP (GA), and TFPO-CSP (ABC) approaches to select the temporal filter parameters (dataset 2)

Iterations	Method	Subject						
		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>
Minimum	TFPO-CSP (PSO)	1	1	1	5	1	1	1
	TFPO-CSP (ABC)	1	1	1	1	1	1	1
	TFPO-CSP (GA)	1	1	1	1	1	1	1
Average	TFPO-CSP (PSO)	1	2	5	14	5	4	1
	TFPO-CSP (ABC)	5	1	2	1	5	5	1
	TFPO-CSP (GA)	3	1	1	1	2	2	1
Maximum	TFPO-CSP (PSO)	4	15	21	22	12	8	7
	TFPO-CSP (ABC)	35	4	8	5	17	21	8
	TFPO-CSP (GA)	25	1	10	2	18	15	1

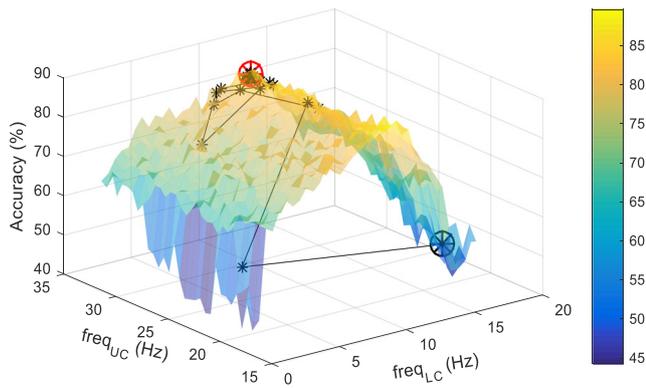


Fig. 5 Surface plot showing accuracies for different frequency band combinations with filter order fixed at 8 using dataset 1 (subject *aa*). An example of the trace of how a chromosome moves through the search space for TFPO-CSP (GA) is also shown with black and red circles indicating start and end points

parameters for each subject are shown in Tables 3 and 4 for dataset 1 and dataset 2, respectively. The minimum number of iterations required by the TFPO-CSP (GA) approach is 1 for all the subjects. The average number of iterations required for selecting the temporal filter parameters using the TFPO-CSP (GA) approach was quite low for all subjects ≤ 3 except for subject *aa* of dataset 1 (having value of 10). For subjects *aa* and *aw* of dataset 1 (using TFPO-CSP (PSO)) and subject *a* of

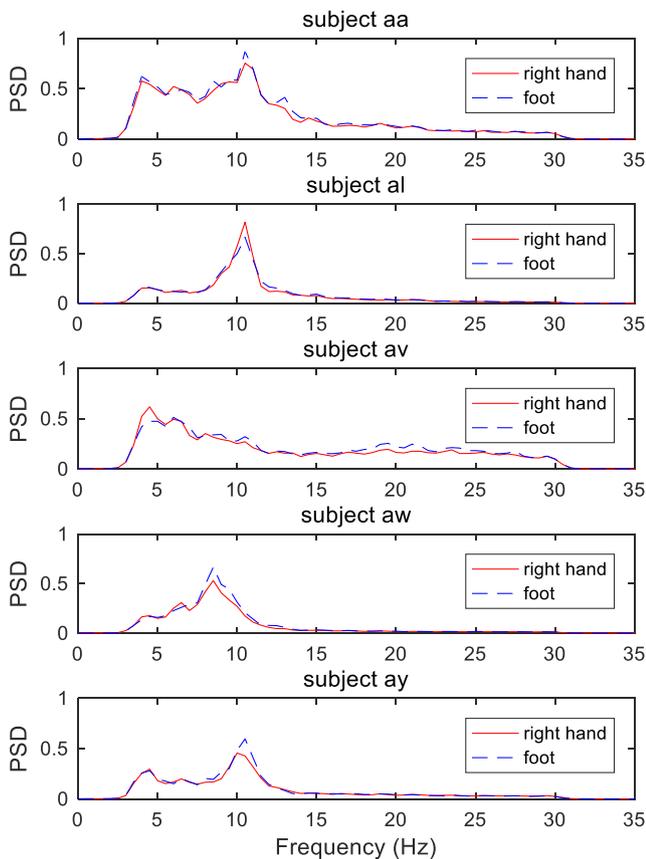


Fig. 6 Average power spectrum density of the EEG dataset for the five subjects of dataset 1 (using all the channels data)

dataset 2 (using TFPO-CSP (ABC)), the maximum number of iterations is the maximum number of iteration set during the experiment (i.e., 35). This may be due to the desired fitness condition set for subjects *aa* and *aw* of dataset 1 and subject *a* of dataset 2 being challenging or due to a local minimum that resulted in the TFPO-CSP (PSO) and TFPO-CSP (ABC), respectively, taking longer time to find the desired solution. TFPO-CSP (GA) obtained the least number of average iterations required to search for the optimal parameters.

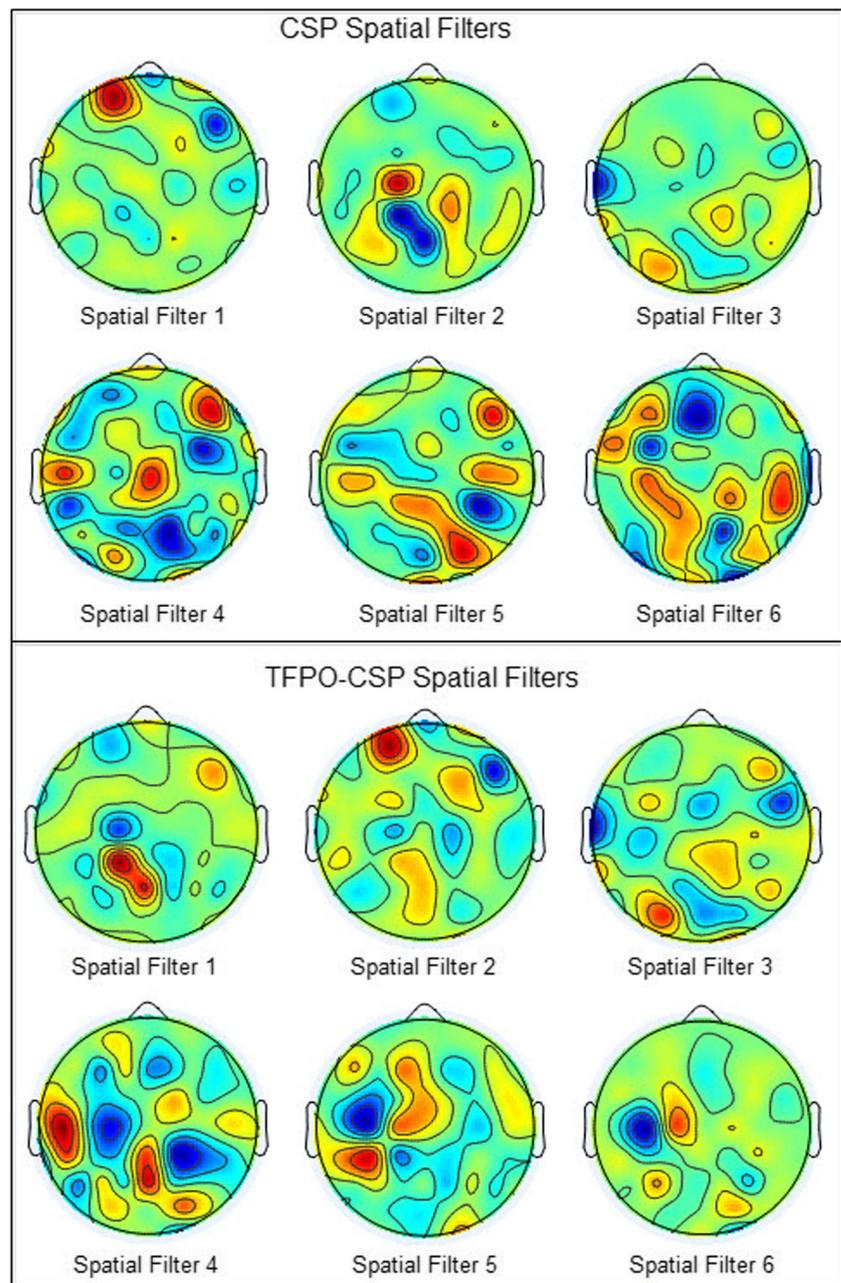
Figure 5 shows the surface plot in 3D space of the classification accuracies (for subject *aa* of dataset 1) when lower and upper cutoff frequencies are varied with a fixed filter order set at 8. It shows that the frequency band selected has a direct impact on the overall performance of the system and justifies the need for searching for filter band parameters. A trace of how a chromosome (for subject *aa* of dataset 1) moves through the search space (for TFPO-CSP (GA)) in order to get to the optimal solution for one of the trial runs is also shown in Fig. 5 with black and red circles indicating the start and end points, respectively. It can be seen from Fig. 5 that the TFPO-CSP (GA) algorithm maneuvers through the search space and successfully finds the optimal solution (shown as connecting lines).

To further verify the performance of our predictor with GA as the optimization algorithm (TFPO-CSP (GA)), the power spectrum density (PSD) of the two-class MI EEG data is conducted. The average PSD of each task for all five subjects derived over trials using all channels data is shown in Fig. 6. It can be noted that the proposed TFPO-CSP (GA) approach effectively searched for a filter band that gives minimum error rate between the two tasks, i.e., the frequency band in which the difference in PSD of the two tasks is largest. This is further explained in Fig. 5 where it is shown that the particle successfully selects the filter parameters that give the optimal solution. Thus, the findings reveal that the most responsive frequency band varies from subject to subject. This explains the need for automatic subject-specific frequency band selection for optimal performance.

4 Discussion

To further show the effectiveness of the proposed method, Fig. 7 shows the topographic plot of the spatial filters that are learned using the CSP and TFPO-CSP methods (spatial filters learned during one of the random runs using subject *aw* of dataset 1). It can be seen that the spatial filters learned by the TFPO-CSP method highly correlate with the two MI tasks (right hand and left foot) as they are contralaterally well localized in the left and right sensorimotor cortex regions. We also tested the classification performance of the learned spatial filters of subject *aw* shown in Fig. 7 and obtained 89.29% and 96.43% classification accuracies (on test data) using

Fig. 7 The learned spatial filters for CSP and TFPO-CSP methods for subject *aw* of dataset 1 (for one of the random trial runs)



CSP and TFPO-CSP spatial filters, respectively. This further validates the results that the spatial filters learned by TFPO-CSP are able to more effectively discriminate between the two MI tasks compared to the spatial filters learned by CSP. This improved performance is due to the use of the most responsive subject-dependent frequency band, which is automatically determined using our proposed TFPO-CSP predictor.

In order to keep the computation complexity of the proposed method to a minimum, we have employed a single frequency band. Thus, the time (test time) taken for a single trial to be processed (temporal filtering, spatial filtering, and feature extraction) and classified using the trained classifier for different methods is shown in Table 5. The time calculated is the test time and does not include the training time as mostly offline training is

Table 5 Test time required by different algorithms for single trial MI EEG signal classification

Method	CSP	CSSP	FBCSP	DFBCSP	SBLFB	TFPO-CSP (GA)
Time (ms)	3.36	12.30	14.22	10.80	13.10	6.96

Table 6 Initialization time and time taken for one iteration during training that is required by different algorithms

Method	TFPO-CSP (PSO)		TFPO-CSP (GA)		TFPO-CSP (ABC)	
	Init	1 iter	Init	1 iter	Init	1 iter
Time (s)	145.44	159.18	128.90	128.95	72.97	145.44

performed for BCI applications. It should be noted that as the number of bands used is increased, the computation time also increases, i.e., computation time is directly proportional to the number of bands used for processing. Our proposed approach uses a single frequency band and thus has a lower computation time compared to that of CSSP, FBCSP, DFBCSP, and SBLFB. The computation time of TFPO-CSP (GA) is higher than that of CSP due to the fact that filter order of 30 (maximum) is used for computing the time for TFPO-CSP (GA) approach. If the filter order is same as that of CSP, then the time for CSP and TFPO-CSP (GA) methods would also be same. The test time for TFPO-CSP (ABC) and TFPO-CSP (PSO) will be same as that of TFPO-CSP (GA) if same filter orders are used. Thus, our approach can be top choice for use with wearable devices, which require algorithms that are computationally less expensive so that it can last longer as they are battery powered. Matlab running on a personal computer at 3.3 GHz (Intel(R) Core(TM) i3) has been used for all processing.

The TFPO-CSP predictor with GA is able to successfully find the optimal filter band parameters that give higher accuracy without the use of fitness condition, i.e., allowing the algorithm to run for maximum number of iterations. However, the drawback is longer training time. Since the training procedure can be performed offline, this drawback can be overcome. Nonetheless, to reduce the training time, fitness condition has been used to converge to the solution in a faster manner. The setting of different fitness conditions for each subject proves vital as the training time is reduced and the proposed TFPO-CSP with GA is still able to achieve higher accuracy. The training time required by TFPO-CSP with PSO, TFPO-CSP with GA, and TFPO-CSP with ABC is given in Table 6, where *Init* is the average initialization time and *1 iter* is the average time required by the respective methods for execution of one complete iteration. Therefore, the overall average training time for each of the methods will be the initialization time plus the time taken for one iteration multiplied by the average number of iterations required by the respective methods. On overall, the TFPO-CSP (GA) approach requires the minimum number of iterations to search for the optimal parameters. Therefore, the TFPO-CSP (GA) approach is recommended over TFPO-CSP (ABC) and TFPO-CSP (PSO).

Our proposed predictor obtained the overall minimum error rates of 10.19% and 19.92% for dataset 1 and dataset 2, respectively. It outperformed all competing methods evaluated

in this work achieving an improvement of 3.54% and 0.65% compared to SBLFB method for dataset 1 and dataset 2, respectively. In comparison with SFBCSP and DFBCSP, our predictor achieved an improvement of 4.22% and 0.75% for dataset 1 while also achieving an improvement of 7.29% and 2.94% for dataset 2, respectively. Our proposed predictor also outperformed the sparsity-aware method [47] (where weighted averaging has been introduced for estimating the average covariance matrix, which is used for calculating the CSP spatial filters) that was evaluated on dataset 1, achieving an improvement of 2.44%. However, the sparsity-aware method performed well compared to CSP method, and can be incorporated in our predictor, which might further improve the performance of the system.

It can be noted that for some of the subjects, the lower cutoff frequencies were in the range of 2–8 Hz, which is lower than the mu rhythm. The proposed method only finds the band that gives optimal performance and guarantees that the selected frequency band contains most of the useful information. However, it does not guarantee that all the frequencies present in the selected frequency band contain useful information. While searching for the optimal parameters, it is possible for the proposed method to find a suitable solution that has relatively lower value for lower cutoff frequency parameter depending on the search space through which the algorithm maneuvers to find the optimal parameters. It is possible that no important information is contained by the signal in the lower frequencies and having or not having these will make no notable difference in the performance. However, it should also be noted that very low values of lower cutoff frequency parameter has been obtained using all the three optimization algorithms for subjects *al* and *ay* of dataset 1 and subject *g* of dataset 2. This suggests that there is some information in the lower frequencies, which is causing all the optimization algorithms used to select the lower value for lower cutoff frequency. Since both datasets have been recorded from healthy subjects, there is very little chance of abnormal behavior. Thus, further research needs to be carried out in order to investigate if these lower frequencies actually contain any useful information about the different MI tasks. However, according to [11], signal in the delta band can be caused by some continuous-attention tasks and signal in theta band can be caused due to a person actively trying to repress an action. Since the trials in both the datasets are cue based and the subjects are required to imagine the MI tasks, it is highly likely that these can be the reason leading to the selection of low values for lower cutoff frequency parameter for some of the subjects.

Furthermore, we have performed paired *t* test with 5% significance level to show the significance of the classification error rate of the proposed method compared to the top 2 performing methods (SBLFB and DFBCSP). The *p* value obtained was 0.045 and 0.034 (compared to SBLFB and DFBCSP, respectively), which shows that significant improvements have been achieved.

The main advantage of the proposed TFPO-CSP predictor is that it automatically finds the responsive frequency band, which contains significant information for maximum discrimination between the MI tasks. This also results in the spatial filter learned to be more effective in discriminating between the different MI tasks. Furthermore, the TFPO-CSP predictor is able to achieve enhanced performance using only a single frequency band compared to the competing methods that utilize multiple sub-bands. This accounts for the low computational complexity of the method and makes it more suitable for implementation on wearable devices. This can help in the development of low-cost BCI systems that will assist people with disabilities to recover their environmental control capabilities. It can also be beneficial in epileptic seizure detection using EEG signals as being able to predict/detect pre-seizure more accurately will save more lives and prevent more accidents. The disadvantage of the TFPO-CSP predictor is that it requires a longer training time compared to other competing methods. However, this problem can be overcome by training the model offline. Moreover, being able to effectively find filter parameters that will produce optimal results in a quicker way is desirable for BCI applications that will require online training. To add on, the range of filter order proposed in this work is specifically for butterworth filter. If other types of filters are to be used, then the users will need to investigate and determine the range for the filter order as different types of filters behave differently. The users can also simply select maximum range up to 50 and the proposed TFPO-CSP predictor will automatically find the best value. However, it might take longer time to find the optimal parameters given a larger search space.

Moreover, the proposed method can be utilized for further tuning the sub-bands selected by other approaches (such as SBCSP, FBCSP, DFBCSP, SFBCSP, SBLFB, and SFTOFSRC) and using dimensionality reduction or feature selection techniques [32–41] to remove redundant features, which can further improve the performance of the existing methods. On the other hand, the use of dictionary optimization together with SRC can further enhance the performance of the TFPO-CSP predictor.

5 Conclusion

In this paper, we proposed a new automatic method of finding the temporal filter parameters for improved MI EEG signal classification. The parameters of a butterworth filter are tuned for optimal performance. The proposed method performed well in selecting subject-specific frequency bands and outperformed state-of-the-art methods achieving the lowest classification error rates. The only drawback of the approach is that it requires a longer training time compared to other methods. However, this drawback can be avoided since

training is usually performed offline for BCI applications. The proposed methods capability in finding the temporal filter parameters that give optimal performance in terms of MI EEG signal classification together with its low computational complexity (for classification of test samples) makes it a strong candidate for future BCI systems.

Acknowledgements Special thanks to the editors and anonymous reviewers for their positive and constructive comments and suggestions that helped improve our manuscript.

Funding information This work was supported by the Faculty Research Committee of the University of the South Pacific, Laboratory for Medical Science Mathematics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan, and College Research Committee of Fiji National University.

References

1. Akram F, Metwally MK, Hee-Sok H, Hyun-Jae J, Tae-Seong K (2013) A novel P300-based BCI system for words typing. In: 2013 International Winter Workshop on Brain-Computer Interface (BCI). pp 24–25. doi: <https://doi.org/10.1109/IWW-BCI.2013.6506617>
2. Andrew C, Pfurtscheller G (1997) On the existence of different alpha band rhythms in the hand area of man. *Neurosci Lett* 222: 103–106. [https://doi.org/10.1016/S0304-3940\(97\)13358-4](https://doi.org/10.1016/S0304-3940(97)13358-4)
3. Ang KK, Chin ZY, Zhang H, Guan C (2008) Filter bank common spatial pattern (FBCSP) in brain-computer interface. In: IEEE International Joint Conference on Neural Networks pp 2390–2397. doi: <https://doi.org/10.1109/IJCNN.2008.4634130>
4. Blankertz B, Dornhege G, Krauledat M, Müller KR, Curio G (2007) The non-invasive Berlin Brain-Computer Interface: fast acquisition of effective performance in untrained subjects. *NeuroImage* 37:539–550. <https://doi.org/10.1016/j.neuroimage.2007.01.051>
5. Cincotti F, Mattia D, Aloise F, Bufalari S, Schalk G, Oriolo G, Cherubini A, Marciani MG, Babiloni F (2008) Non-invasive brain-computer interface system: towards its application as assistive technology. *Brain Res Bull* 75:796–803. <https://doi.org/10.1016/j.brainresbull.2008.01.007>
6. Dornhege G, Blankertz B, Curio G, Müller K (2004) Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. *IEEE Trans Biomed Eng* 51: 993–1002. <https://doi.org/10.1109/TBME.2004.827088>
7. Dornhege G, Blankertz B, Krauledat M, Losch F, Curio G, Müller KR (2006) Combined optimization of spatial and temporal filters for improving brain-computer interfacing. *IEEE Trans Biomed Eng* 53:2274–2281. <https://doi.org/10.1109/TBME.2006.883649>
8. Higashi H, Tanaka T (2013) Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification. *IEEE Trans Biomed Eng* 60:1100–1110. <https://doi.org/10.1109/TBME.2012.2215960>
9. Janjarasjitt S (2017) Epileptic seizure classifications of single-channel scalp EEG data using wavelet-based features and SVM. *Med Biol Eng Comput* 55:1743–1761. <https://doi.org/10.1007/s11517-017-1613-2>
10. Karaboga D (2005) An idea based on honey bee swarm for numerical optimization. Technical Report TR06: Erciyes University
11. Kimmizi-Alsan E, Bayraktaroglu Z, Gurvit H, Keskin YH, Emre M, Demiralp T (2006) Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological

- markers of response inhibition and sustained attention. *Brain Res* 1104:114–128. <https://doi.org/10.1016/j.brainres.2006.03.010>
12. Kleih SC, Kuafmann T, Zickler C et al (2011) Out of the frying pan into the fire—the P300-based BCI faces real-world challenges. *Prog Brain Res* 194:27–46. <https://doi.org/10.1016/B978-0-444-53815-4.00019-4>
 13. Koles ZJ (1991) The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. *Electroencephalogr Clin Neurophysiol* 79:440–447. [https://doi.org/10.1016/0013-4694\(91\)90163-X](https://doi.org/10.1016/0013-4694(91)90163-X)
 14. Kumar S, Mamun K, Sharma A (2017) CSP-TSM: optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI. *Comput Biol Med* 91:231–242. <https://doi.org/10.1016/j.combiomed.2017.10.025>
 15. Kumar S, Sharma A, Mamun K, Tsunoda T (2016) A Deep learning approach for motor imagery EEG signal classification. In: 3rd Asia-Pacific World Congress on Computer Science and Engineering. doi: <https://doi.org/10.1109/APWC-on-CSE.2016.017>
 16. Kumar S, Sharma A, Tsunoda T (2017) An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information. *BMC Bioinformatics* 18: 545. <https://doi.org/10.1186/s12859-017-1964-6>
 17. Kumar S, Sharma R, Sharma A, Tsunoda T (2016) Decimation filter with common spatial pattern and fishers discriminant analysis for motor imagery classification. In: IEEE World Congress on Computational Intelligence doi: <https://doi.org/10.1109/IJCNN.2016.7727457>
 18. La Rocca D, Campisi P, Sole-Casals J (2013) EEG based user recognition using BUMP modelling. In: International Conference of the Biometrics Special Interest Group (BIOSIG). pp 1–12
 19. Lemm S, Blankertz B, Curio G, Muller K (2005) Spatio-spectral filters for improving the classification of single trial EEG. *IEEE Trans Biomed Eng* 52:1541–1548. <https://doi.org/10.1109/TBME.2005.851521>
 20. McFarland DJ, Sarnacki WA, Wolpaw JR (2010) Electroencephalographic (EEG) control of three-dimensional movement. *J Neural Eng* 7:036007. <https://doi.org/10.1088/1741-2560/7/3/036007>
 21. Miao M, Wang A, Liu F (2017) A spatial-frequency-temporal optimized feature sparse representation-based classification method for motor imagery EEG pattern recognition. *Med Biol Eng Comput* 55: 1589–1603. <https://doi.org/10.1007/s11517-017-1622-1>
 22. Mumtaz W, Ali SSA, Yasin MAM, Malik AS (2018) A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD). *Medi Biol Eng Comput* 56:233–246. <https://doi.org/10.1007/s11517-017-1685-z>
 23. Naveen, R.S., and Julian, A. (2013). Brain computing interface for wheel chair control. In: Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT). pp 1–5. doi: <https://doi.org/10.1109/ICCCNT.2013.6726572>
 24. Nicolas-Alonso LF, Gomez-Gil J (2012) Brain computer interfaces, a review. *Sensors* 12:1211–1279. <https://doi.org/10.3390/s120201211>
 25. Novi Q, Cuntai G, Dat TH, Ping X (2007) Sub-band common spatial pattern (SBCSP) for brain-computer interface. In: 3rd International IEEE/EMBS Conference on Neural Engineering. pp 204–207. doi: <https://doi.org/10.1109/CNE.2007.369647>
 26. Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol* 110:1842–1857. [https://doi.org/10.1016/S1388-2457\(99\)00141-8](https://doi.org/10.1016/S1388-2457(99)00141-8)
 27. Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. *Swarm Intelligence* 1:33–57
 28. Ramesh S, Krishna MG, Nakirekanti M (2014) Brain computer interface system for mind controlled robot using Bluetooth. *Int J Comput Appl* 104:20–23
 29. Ramoser H, Muller-Gerking J, Pfurtscheller G (2000) Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans Rehabil Eng* 8:441–446. <https://doi.org/10.1109/86.895946>
 30. Samiee K, Kovcs P, Gabbouj M (2017) Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local Gabor binary patterns feature extraction. *Knowledge Based Syst* 118:228–240. <https://doi.org/10.1016/j.knosys.2016.11.023>
 31. Serruya MD (2014) Bottlenecks to clinical translation of direct brain-computer interfaces. *Front Syst Neurosci* 8:226. <https://doi.org/10.3389/fnsys.2014.00226>
 32. Sharma A, Imoto S, Miyano S (2012) A between-class overlapping filter-based method for transcriptome data analysis. *J Bioinforma Comput Biol* 10:1250010. <https://doi.org/10.1142/S0219720012500102>
 33. Sharma A, Imoto S, Miyano S (2012) A filter based feature selection algorithm using null space of covariance matrix for DNA microarray gene expression data. *Current Bioinform* 7:289–294. <https://doi.org/10.2174/157489312802460802>
 34. Sharma A, Imoto S, Miyano S (2012) A top-r feature selection algorithm for microarray gene expression data. *IEEE/ACM Trans Comput Biol Bioinform* 9:754–764. <https://doi.org/10.1109/TCBB.2011.151>
 35. Sharma A, Imoto S, Miyano S, Sharma V (2012) Null space based feature selection method for gene expression data. *Int J Mach Learn Cybernetics* 3:269–276. <https://doi.org/10.1007/s13042-011-0061-9>
 36. Sharma A, Koh C, Imoto S, Miyano S (2011) Strategy of finding optimal number of features on gene expression data. *Electron Lett* 47:480–482. <https://doi.org/10.1049/el.2011.0526>
 37. Sharma A, Paliwal KK (2015) A deterministic approach to regularized linear discriminant analysis. *Neurocomputing* 151:207–214. <https://doi.org/10.1016/j.neucom.2014.09.051>
 38. Sharma A, Paliwal KK (2007) Fast principal component analysis using fixed-point algorithm. *Pattern Recogn Lett* 28:1151–1155. <https://doi.org/10.1016/j.patrec.2007.01.012>
 39. Sharma A, Paliwal KK (2015) Linear discriminant analysis for the small sample size problem: an overview. *Int J Mach Learn Cybernetics* 6:443–454. <https://doi.org/10.1007/s13042-013-0226-9>
 40. Sharma A, Paliwal KK (2012) A new perspective to null linear discriminant analysis method and its fast implementation using random matrix multiplication with scatter matrices. *Pattern Recogn* 45: 2205–2213. <https://doi.org/10.1016/j.patcog.2011.11.018>
 41. Sharma A, Paliwal KK, Imoto S, Miyano S (2013) Principal component analysis using QR decomposition. *Int J Mach Learn Cybernetics* 4:679–683. <https://doi.org/10.1007/s13042-012-0131-7>
 42. Shin Y, Lee S, Ahn M, Cho H, Jun SC, Lee H-N (2015) Simple adaptive sparse representation based classification schemes for EEG based brain-computer interface applications. *Comput Biol Med* 66:29–38. <https://doi.org/10.1016/j.combiomed.2015.08.017>
 43. Silvoni S, Ramos-Murguialday A, Cavinato M, Volpato C, Cisetto G, Turlora A, Piccione F, Birbaumer N (2011) Brain-computer interface in stroke: a review of progress. *Clin EEG Neurosci* 42:245–252. <https://doi.org/10.1177/155005941104200410>
 44. Sohrabpour A, Lu Y, Kankirawatana P, Blount J, Kim H, He B (2015) Effect of EEG electrode number on epileptic source localization in pediatric patients. *Clin Neurophysiol* 126:472–480. <https://doi.org/10.1016/j.clinph.2014.05.038>
 45. Suk HI, Lee SW (2013) A novel Bayesian framework for discriminative feature extraction in brain-computer interfaces. *IEEE Trans Pattern Anal Mach Intell* 35:286–299. <https://doi.org/10.1109/TPAMI.2012.69>

46. Thomas KP, Cuntai G, Lau CT, Vinod AP, Keng AK (2009) A new discriminative common spatial pattern method for motor imagery brain computer interfaces. *IEEE Trans Biomed Eng* 56:2730–2733. <https://doi.org/10.1109/TBME.2009.2026181>
47. Tomida N, Tanaka T, Ono S, Yamagishi M, Higashi H (2015) Active data selection for motor imagery EEG classification. *IEEE Trans Biomed Eng* 62:458–467. <https://doi.org/10.1109/TBME.2014.2358536>
48. Wang H, Zhang Y, Waytowich NR, Krusienski DJ, Zhou G, Jin J, Wang X, Cichocki A (2016) Discriminative feature extraction via multivariate linear regression for SSVEP-based BCI. *IEEE Trans Neural Syst Rehabil Eng* 24:532–541. <https://doi.org/10.1109/TNSRE.2016.2519350>
49. Wei Q, Wei Z (2015) Binary particle swarm optimization for frequency band selection in motor imagery based brain-computer interfaces. *Bio-Med Mater Eng* 26:S1523–S1532. <https://doi.org/10.3233/BME-151451>
50. Widmann A, Schröger E (2012) Filter effects and filter artifacts in the analysis of electrophysiological data. *Front Psychol* 3. <https://doi.org/10.3389/fpsyg.2012.00233>
51. Woehrle H, Krell MM, Straube S, Kim SK, Kirchner EA, Kirchner F (2015) An adaptive spatial filter for user-independent single trial detection of event-related potentials. *IEEE Trans Biomed Eng* 62:1696–1705. <https://doi.org/10.1109/TBME.2015.2402252>
52. Xu Y, Fan P, Yuan L (2013) A simple and efficient artificial bee colony algorithm. *Mathematical Probl Eng* 9. doi: <https://doi.org/10.1155/2013/526315>
53. Yu T, Xiao J, Wang F, Zhang R, Gu Z, Cichocki A, Li Y (2015) Enhanced motor imagery training using a hybrid BCI with feedback. *IEEE Trans Biomed Eng* 62:1706–1717. <https://doi.org/10.1109/TBME.2015.2402283>
54. Zhang Y, Wang Y, Jin J, Wang X (2017) Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. *Int J Neural Syst* 27:1650032. <https://doi.org/10.1142/S0129065716500325>
55. Zhang Y, Wang Y, Zhou G, Jin J, Wang B, Wang X, Cichocki A (2018) Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces. *Expert Syst Appl* 96:302–310. <https://doi.org/10.1016/j.eswa.2017.12.015>
56. Zhang Y, Zhou G, Jin J, Wang X, Cichocki A (2015) Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface. *J Neurosci Methods* 255:85–91. <https://doi.org/10.1016/j.jneumeth.2015.08.004>
57. Zhang Y, Zhou G, Jin J, Zhao Q, Wang X, Cichocki A (2016) Sparse Bayesian classification of EEG for brain computer interface. *IEEE Trans Neural Netw Learn syst* 27:2256–2267. <https://doi.org/10.1109/TNNLS.2015.2476656>
58. Zhou G, Zhao Q, Zhang Y, Adali T, Xie S, Cichocki A (2016) Linked component analysis from matrices to high-order tensors: applications to biomedical data. *Proc IEEE* 104:310–331. <https://doi.org/10.1109/JPROC.2015.2474704>



Shiu Kumar is currently pursuing his Ph.D. degree at School of Engineering and Physics, The University of the South Pacific, Fiji. His research interests include brain-computer interface, bio-medical signal processing, and machine learning.



Alok Sharma received the Ph.D. degree in the area of pattern recognition from Griffith University, Brisbane, Australia, in 2006. He is a Scientist at RIKEN Center for Integrative Medical Sciences, Japan. He is also an Adjunct Associate Professor at the Institute for Integrated and Intelligent Systems (IIIS), Griffith University, Australia. He is concurrently serving as a Professor at USP. His research interests include pattern

recognition, human cancer classification, and proteomics. Website: www.alok-ai-lab.com