



Optimization of parameters for improving the performance of EEG-based BCI system

Mandeep Kaur Ghumman¹ · Satvir Singh¹ · Navtej Singh² · Balkrishan Jindal³

Received: 24 January 2020 / Accepted: 7 October 2020 / Published online: 22 October 2020
© Springer Nature Switzerland AG 2020

Abstract

Brain–computer interface (BCI) is an active domain which has attracted attention of the research community in recent years. It offers huge potential as a technology which can estimate the intention of a user by analysis of brain signals and establish a communication channel directly between a human brain and an external device. Electroencephalography (EEG) is the most popular signal acquisition technique due to its ease of use and simplicity. In EEG-based BCI systems, electrodes are placed on specific positions on the scalp of the subject to record electrical activity. The BCI system consists of sequential stages of signal acquisition, its preprocessing, feature extraction and feature classification. It is an active research area which has a focus on improving classification accuracy in motor imagery-based BCI systems. The first stage in a BCI system is to acquire EEG signals from different positions of the scalp of the human subject. The acquired brain signals are preprocessed to remove artifacts before these are fed to feature the extraction stage. In this paper, independent component analysis (ICA) technique is used to remove artifacts from acquired signals. Filter bank common spatial pattern (FBCSP) technique is then used for feature extraction and feature selection. A feature classification approach based on support vector machine (SVM) is proposed in this work and its performance is enhanced by optimizing its polynomial kernel parameters. Selection of kernel parameters is done by grid search method using the fivefold cross-validation procedure. The proposed approach is then executed on publicly available data set 2a of BCI Competition IV. Results show that the proposed approach offers higher classification accuracy and lower misclassification rate as compared to other methods executed on the same dataset, as reported in literature.

Keywords Brain–computer interface · EEG · Motor imagery · Support vector machine · Polynomial kernel · Filter bank common spatial pattern · Independent component analysis

1 Introduction

The brain–computer interface (BCI) is a technique of establishing a channel for communication between the brain of the user and an external device without using the brain's normal nerve pathways to other body parts [1]. It provides an advanced technology which can translate the intention of a user from the analysis of brain signals directly into corresponding commands to establish a communication channel directly between the human brain and external devices [2]. BCI is a multi-disciplinary research field involving neuro-

logy, rehabilitation engineering, human–computer interaction (HCI), signal processing and machine learning [3]. In recent years, the brain signals have been extensively analyzed and explored for BCI applications. Electroencephalography (EEG) is an important tool for recording functional brain activity. It is the most used signal acquisition technique for MI-based BCIs due to its simplicity and ease of use [4]. It offers better temporal resolution at a lower cost, which makes it popular among researchers [5]. Professor Hans Berger from Germany discovered in 1924 that electrical signals produced by the human brain could be recorded from the scalp using electrodes [6]. He developed the technique of electroencephalography (EEG) for fetching electrical signals of the brain. Although EEG is a popularly used mechanism for fetching brain signals, other techniques like magnetoencephalography (MEG) and electrocorticography (ECoG) can also be used to monitor the activities of the human brain. The availability of powerful computer equipment at lower

✉ Mandeep Kaur Ghumman
er.mandeepghumman@gmail.com

¹ IKG Punjab Technical University, Kapurthala, Punjab, India

² Punjabi University, Patiala, Punjab, India

³ YCoE, Punjabi University, Guru Kashi Campus, Talwandi Sabo, Punjab, India

costs and new insights into the functionality of the human brain has encouraged researchers to focus on developing new supplementary communication and control technology for patients suffering from neuromuscular disorders due to sclerosis, brain stroke, spinal cord injury, etc. [7]. The reliability of a BCI system for rehabilitation of such patients is of paramount importance. The safety of such systems can be enhanced by improvement in hardware as well as using advanced techniques such as machine learning to make such system intelligent and reliable.

BCIs can be used in a variety of areas including prosthetic limbs, mobility devices, robotics and device communication. These developments have led to improvements in the techniques of processing of signals emanating from the scalp during the performance of a specific type of mental task. The major objective of BCI research is to develop supplementary systems that allow disabled users to control artificial limbs and communicate with the outer environment. The electrodes are placed as per standardized international 10–20 system of EEG, on different parts of the scalp of the human subject to record the electrical activity of the brain [8]. These signals acquired from electrodes reflect the motor imagery (MI) activity of the subject, such as the hand, foot or tongue movements. [9]. For the operation of every BCI, a neurological control signal is required. Different BCI systems have been developed on the basis of this control signal, [10]. Most of the current BCI systems using EEG fetch their input from neurological phenomena such as P300 potential, event-related potential (ERP), Mu and/or beta rhythms with event-related synchronization (ERS) and event-related desynchronization (ERD), cognitive task-related EEGs, visual evoked potential (VEP) and slow cortical potentials (SCP). Subjects using the BCI induce brain activity pattern, by following the experimental protocol for that particular BCI approach. The protocol followed by the subjects can be MI movements and focusing on visual clue of flashing characters on screen. Motor imagery (MI) is a common paradigm used in BCI. It is an MI task in which a subject is cued to just imagine the movement of a specific limb, without actually executing its action. EEG signals are then recorded while executing multiple MI tasks of hands, feet, tongue movement, etc. [11]. Movement or even preparation for movement leads to a decrease in mu and beta rhythms, which is called ERD. In contrast to this, mu and beta rhythm increase after the movement is completed, which is called ERS. The ERD and ERS do not require the actual movement, but occur with mere imagination of movement, which is called motor imagery (MI), and hence can effectively be used for BCI. The efficiency of a BCI system depends on the choice of suitable algorithms for implementation of its various stages [12]. It is important to choose a suitable classifier from the range of well-known classifiers such as linear discriminant analysis (LDA), support vector machine (SVM), fuzzy logic (FL), k-nearest neigh-

bor algorithm (KNN) and artificial neural network (ANN) for EEG classification [13]. For multi-class classification, selection of a particular classifier is a critical issue in the BCI system [14]. Linear classifiers are generally preferred for EEG classification due to their low computational complexity and better stability [15]. They are also less prone to overfitting problem as compared to nonlinear classifiers, especially when only a limited number of samples are available [16]. The main objective of the BCI-based applications is to accurately translate the brain wave patterns extracted from the EEG signals into the desired machine commands. The objective of many researches is to enhance this accuracy of interpretation of the harvested EEG signals [17]. SVM is a popularly used classifier for MI-based BCI systems. It establishes an optimum hyperplane to separate different classes as far as possible [18]. It can implement multi-class classification and is immune to curse-of-dimensionality of data. Selection of suitable kernel parameters in SVM is of paramount importance to obtain classification accuracy results [19]. Parameters value can be varied to set the the boundary decision in the classifier [20]. This work presents selection of suitable kernel and setting optimal values of the kernel parameters to obtain the decision function, which enhances classification accuracy and overall performance of the MI-based BCI system. In this work, SVM with polynomial kernel (SVM-PK) approach is proposed for EEG signal classification in MI-based BCI system. The performance is improved by selecting the optimal values of the polynomial kernel, by using the grid-search method. These values are then varied to obtain better classification accuracy, which is evaluated by using K-fold cross-validation procedure. This work has improved the performance of MI-based BCI system by enhancing the classification accuracy of MI data, which is then compared with other methods executed on the same dataset, as reported in literature [21].

2 Related work

The goal of the classification stage in BCI is to automatically assign a class to the feature vector, which was extracted in the previous stage. It represents the mental task performed by the BCI user. Classification is obtained by executing algorithms called classifiers. The researchers have explored different methods to implement classifiers for identification of the class to which the feature vector belongs.

Garrett et al. [22] reported the results of implementation of one linear (LDA) and two nonlinear classifiers (NN and SVM) for classification of spontaneous EEG signals, while subjects performed five mental tasks. They concluded that SVMs provide a powerful method for data classification as it uses machine learning and artificial intelligence (AI), for systematic exploration of the EEG feature classification.

Kamoussi et al. [23] proposed a novel approach of using source analysis for classifying MI tasks. They proposed two-equivalent-dipoles analysis for classification of signals recorded from 15 channels from sensory motor area of four subjects. They used noise normalization, spatial filtering, time–frequency analysis and independent component analysis for preprocessing of these signals and reported 80% classification accuracy.

Pfurtscheller et al. [24] reported that phase information and adaptive classification can improve the performance of a BCI and also reduce its training time. They reported that by use of high harmonics features for classification, the performance of a four-class BCI system can be improved. They also demonstrated that feedback can modify sensory motor rhythms and recommended powerful algorithms to search for electrode placement locations.

Bhuvaneswari et al. [18] reviewed different kernel functions in SVM, which is a machine learning method for classification of EEG signals in MI-based BCIs. They used ICA for preprocessing and removing the artifacts, to improve signal-to-noise ratio. They discussed the important role of kernel function in nonlinear separable methods, while using SVM for classification of EEG signals.

Ilyas et al. [25] reviewed the selection of appropriate algorithms for preprocessing, feature extraction and feature classification in a BCI system. They have discussed their advantages, disadvantages and current trends of BCI research.

Mahmood et al. [16] considered mu and beta frequency ranges of recorded EEG signals for MI-based BCI system. They employed CSP for feature extraction and SVM for classification of these signals. They evaluated their approach on dataset IIIa of BCI competition III, and observed improvement in classification accuracy for online BCI systems.

Arbabi et al. [26] compared the effect of different types of selected features and classification algorithms for classifying brain signals in MI-based BCI systems. The results showed that statistical features and signal energy in different frequency bands are among the most appropriate features, which can be processed for implementing a BCI system.

Zhang et al. [27] introduced both cascade and parallel convolutional recurrent neural network models for estimating the intended movements by analysis of raw EEG data. They evaluated their performance on a large-scale movement intention EEG dataset fetched from 108 subjects, and investigated the influence of the spatio-temporal information on the performance of the proposed BCI system.

Lotte et al. [9] surveyed existing literature and reported that there is a need for validation of techniques on off-line as well online BCI systems. They emphasized that calibration of such systems should improve their convenience and robustness against real-life noise in EEG signals. They suggested that the techniques used in BCI systems should be

invariant over time, users and contexts. They recommended the use of new generation of BCI classification methods that process human feedback, so that they can adapt to user states, traits and skills.

The authors in [21] have proposed a fuzzy logic system (FLS)-based approach for multi-class MI data classification. They fused the fuzzy system with particle swarm optimization (PSO) method for improving the classification performance. They used CSP algorithm in the feature extraction phase to extract relevant discriminant features from multi-class EEG data. The learning process of an FLS is computationally intensive. Hence, they reduced the computational expense of the multi-class FLS-based BCI system by application of PSO to reduce processing time. They cross-validated the performance of the proposed FLS method on benchmark data sets, and suggested studying more efficient feature extraction and selection methods in future research to improve the classification performance of a BCI system.

The author in [28] has presented a deep learning approach for classification of MI-based BCI using an adaptive method to determine the threshold. The widely used common spatial pattern (CSP) method is used to extract the variance-based CSP features, which is then fed to the deep neural network for classification. They presented a framework for use of deep neural network (DNN) for MI-BCI classification and evaluated the effectiveness of the proposed framework on dataset IVa of the BCI Competition III.

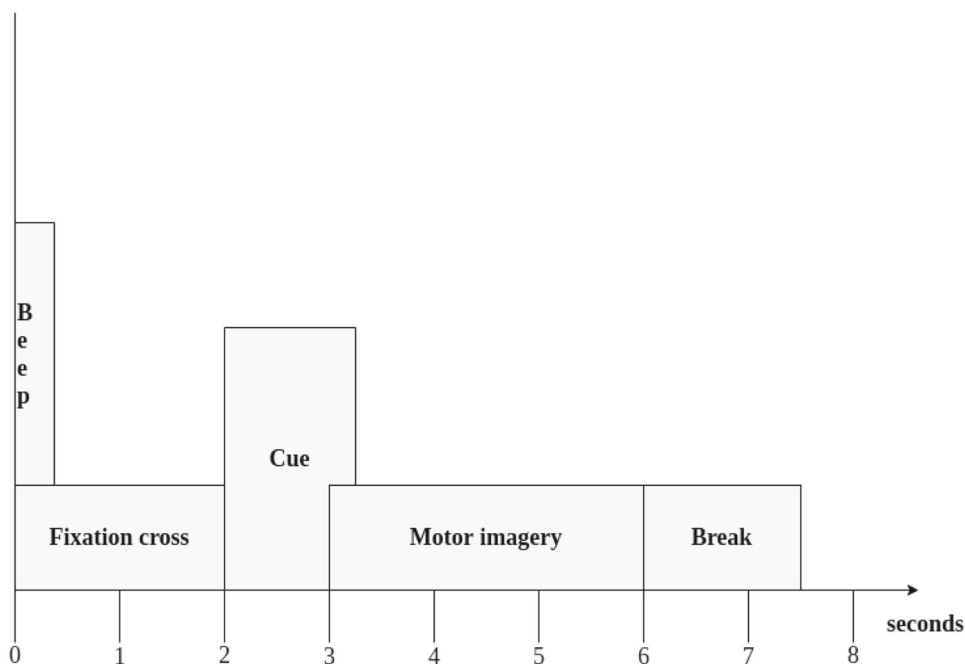
3 Dataset and methodology

The MI-based BCI paradigm is widely used in a variety of applications. It has shown better potential for rehabilitation of patients suffering from motor impairments. It can provide them with an alternative mechanism to communicate with the external world just by thinking of a motor task, without actually performing the movement. In this work, publicly available benchmark EEG dataset from BCI competition IV is used, to evaluate the classification accuracy of our proposed approach [29]. This dataset is extensively used by the BCI research community and contains four-class MI data which is described in the next section

3.1 Experimental paradigm

In this work, the data set 2a from the BCI competition IV [29] is used. This data set is publicly available for the research community and consists of EEG data recorded from nine subjects. Two sessions were recorded for a given subject on two different days. Each session consisted of six runs and each run consisted of 48 trials (12 trials for each motor imagery class). In each trial, a cue was shown on the screen

Fig. 1 Timing of the dataset 2a from BCI competition IV [31]



instructing the subject to perform one of the four MI tasks using the left hand, right hand, both feet, or tongue movement [30].

A single session during the experiment consisted of 288 trials, 72 trials for each of the four MI tasks. Each trial started with a short sound (warning tone) and a fixation cross was shown on the computer screen. After 2 s, in the place of the fixation cross, a cue was shown (in a form of a small arrow) telling the subject to start the corresponding motor imagery task. After another 1.25 s, the arrow changed back to the fixation cross. The MI task is continued until the sixth second of the trial when the fixation cross disappeared. After that there was a short break where the screen was black again. The timing diagram of data acquisition is shown in Fig. 1.

4 Preprocessing to remove artifacts from EEG signals

The acquired data contain a lot of noise from external sources, hence it is necessary to remove these artifacts from EEG signals, in the preprocessing stage, which influences the performance of the overall BCI system [3]. The recorded data contain undesired signals such as electrooculography (EOG), electromyography (EMG), electrocardiography (ECG) and power line noise signals. The methods used for preprocessing depends on the noise levels present in raw signals as well as on techniques used in further processing of the data.

Simple frequency-specific filtering techniques are not sufficient to remove these noise signals due to their overlapping

spectral characteristics and poor spatial resolution of EEG signals. Hence, sophisticated spatial filtering methods such common spatial pattern (CSP), principal component analysis (PCA) and independent component analysis (ICA) are popularly used in the preprocessing stage to reduce these noise signals and improve its spatial resolution. This stage aims at cleaning and denoising the recorded digital data for enhancing the relevant information embedded in the signals.

In this study, ICA is used to remove artifacts from EEG signals and isolate the required information from these signals. It is a computational method which separates signals from multiple sources into subcategories, based on their statistical independence [32]. It performs the separation of independent components by maximizing their statistical independence. ICA is applied to remove EOG, EMG and ECG artifacts from the acquired signals. In the dataset used in this study, there are 22 EEG channels and 3 EOG channels for recording the signals from electrodes placed on the scalp of the subject. ICA is used to remove three EOG channels related to the movement of the eyes [33].

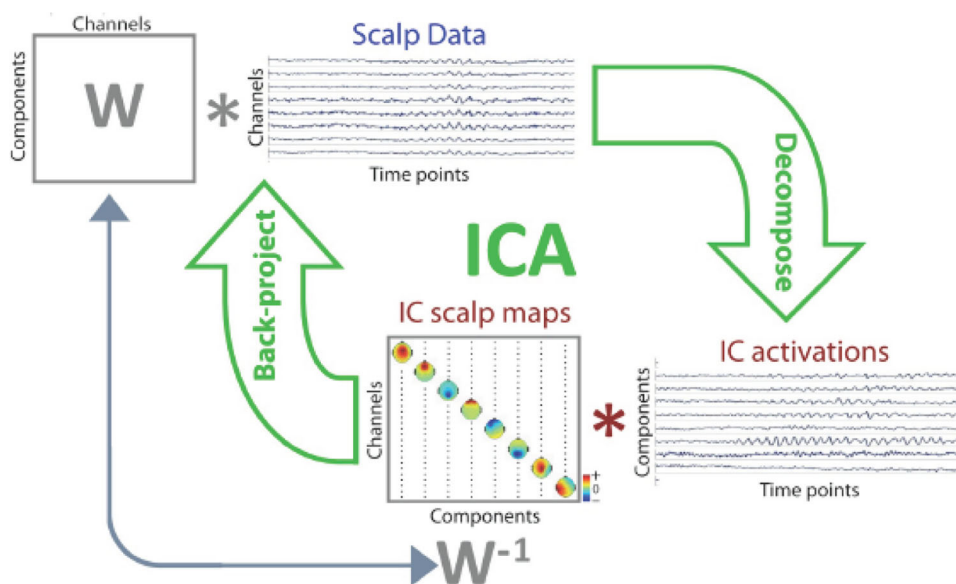
4.1 Independent component analysis

ICA is often used for detection and removal of the eye, muscle, and line noise artifacts.

The EEG activity observed at different electrodes placed on the scalp overlaps and generates some redundant information. ICA is used to separate the artifacts acquired from multiple electrodes.

ICA application to a matrix of EEG scalp data finds an 'unmixing' matrix of weights (W). This matrix is then

Fig. 2 Flowchart for ICA data decomposition and back projection



multiplied by the scalp data matrix to generate a matrix of independent component (IC) activities, as shown in Fig. 2. EEGLAB toolkit is used in this study, for an automated version of the infomax ICA algorithm.

4.2 Feature extraction

The successful classification of MI tasks can be achieved by the successful extraction of the required features from the EEG signals. CSP is a popularly used feature extraction method for the MI-based BCI system. It establishes linear subspaces, so that the variance value of one of the projected class is enhanced to a maximum degree. Simultaneously, the variance value of the other class is reduced to a minimum. The optimal identified spatial filters are established by collective diagonalization of the two covariance matrices, which is calculated from two classes of the EEG signals. In the first step, the normalized covariance matrix of the fetched EEG signal *E* in each trial is calculated as:

$$R = EE' / (\text{trace}(EE')), \tag{1}$$

where *E* denotes *n* × *t* matrix, *n* is the number of channels and *t* is the number of samples. The average of covariance matrices calculated from trials within a class, *M_a* and *M_b*, is added to produce a comprehensive covariance matrix *M_c* = *M_a* + *M_b*. The eigenvectors *E_c* and eigenvalues λ of this covariance matrix result in whitening transform

$$W = \lambda^{-1/2} E'_c, \tag{2}$$

where *M_c* = *E_c*λ*E'_c*. Then, *M_a* and *M_b* are transformed by

$$S_a = WM_aW', S_b = WM_bW'. \tag{3}$$

The values of *S_a* and *S_b* are calculated by using the same eigenvectors, in such a way that *S_a* = *U*ψ_{*a*}*U'* and *S_b* = *U*ψ_{*b*}*U'*. *U* is calculated from the common orthonormal eigenvectors of *S_a* and *S_b* and the values of ψ_{*a*} and ψ_{*b*} represent the calculated diagonal matrices of eigenvalues, such that they add up to a total of 1.

Subsequently, both classes a and b are projected onto *U*₁, which is the first eigenvector, leading to class a yielding the maximal value of variance and class b yielding the minimal value of variance. In contrast to this, when both classes are computationally projected onto the last of the eigenvector *U_m*, then the class a attains the minimal value of variance, whereas the class b attains the maximal value of variance. In implementation, only a few of the eigenvectors are selected, *U** = *U*₁, . . . , *U_m*, *U_{N-m+1}*, . . . , *U_N*, wherein the value of *m* is low (*m* ≪ *N*). The finally calculated projection matrix is represented as

$$P = U * W. \tag{4}$$

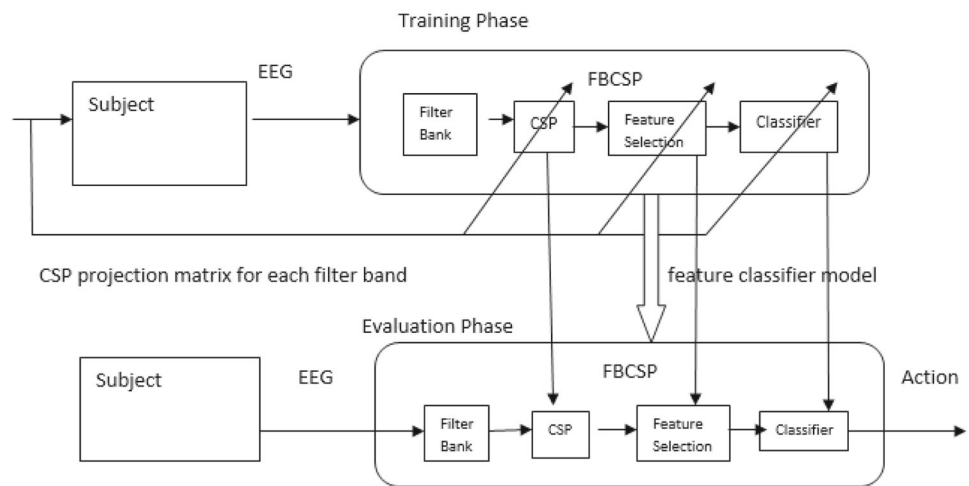
The attained dimension of the original signals is reduced to '2*m*' as per the following equation:

$$f_p = \log \left(\frac{\text{var}(Z_p)}{\sum_{p=1}^{2m} \text{var}(z_p)} \right), p = 1, \dots, 2m \tag{5}$$

The logarithmic transformation method is used to extract normal distributed elements in *f*.

This work has focused on the motor and sensorimotor rhythms, which fall in the frequency band of 8–30 Hz, by using band pass filter to select the relevant band. The patterns of MI signals are distinguished by temporal/spectral and spatial filters. Hence, their optimization has a direct impact

Fig. 3 Filter bank common spatial pattern [34]



on the performance of a BCI system. Feature extraction is implemented by using FBCSP [34].

The FBCSP algorithm is illustrated in Fig. 3. It has stages of signal processing and then implementation of machine learning procedure on the processed EEG data. It first filters EEG signals in multiple frequency bands using a filter bank. It uses band pass filters in having a range of 0.5–40 Hz. Spatial filters are optimized, for each of the filter band, using the classic CSP algorithm. Finally, among the multiple spatial filters obtained, the best resulting features are selected using feature selection algorithms using mutual information-based feature (MIBIF) selection. The MIBIF method selects both the best spectral and spatial filters as each feature corresponds to a single frequency band and CSP spatial filter. It calculates mutual information for each of the feature and arranges it in decreasing order. The top k features are selected from this list for the next stage of classification.

4.3 Classification

Various classification algorithms can be used for a BCI system. The choice of such a classification algorithm depends on many factors including the BCI paradigm used and type of recorded input data. The efficiency of the classifier has a critical effect on the performance and accuracy attained by the BCI system. In this work, linear classifiers are analyzed. These types of classifiers use linear functions to demarcate different MI classes. LDA and SVM are two main linear classifiers used in MI-based BCIs. The LDA technique has a very low computational requirement which makes it suitable for an online BCI system. SVM is efficient for synchronous BCI due to its regularization property and immunity to the curse-of-dimensionality problem. In this paper, the performance of classifiers is analyzed and compared on dataset 2a of BCI Competition IV [15].

4.4 Support vector machine

Support vector machine is a popular classifier for MI-based BCI systems to classify EEG signals. It establishes a hyper-plane separating all data points belonging to one class from the ones belonging to other classes. It creates decision boundaries by using support vectors. It separates different classes by mapping the data to a higher-dimensional space. It tries to maximize the margins by using a kernel function [36]. It allows parameter adjustment to improve the classification rate. The values of the degree of the kernel and regularization parameter, represented as d and C , respectively, are chosen to adjust the balance between algorithmic complexity and number of non-separable points.

SVM increases the margin of separation between various classes and tries to reduce the classification error to a minimum, for data points represented by the respective slack variables, as represented in Fig. 4. The SVM for a k -class problem with n -training points can be represented as a min-

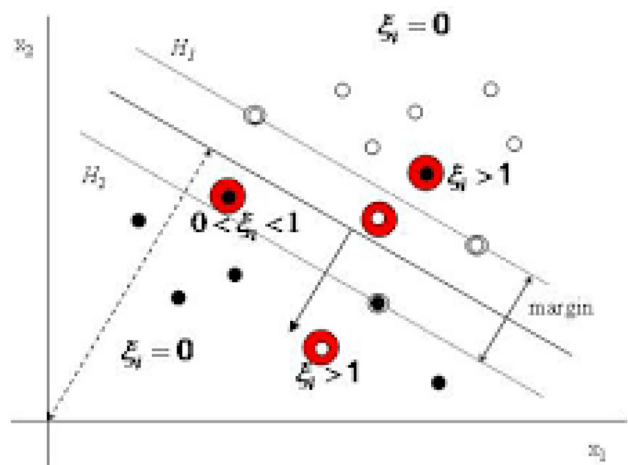


Fig. 4 Separation of nonlinear data points [35]

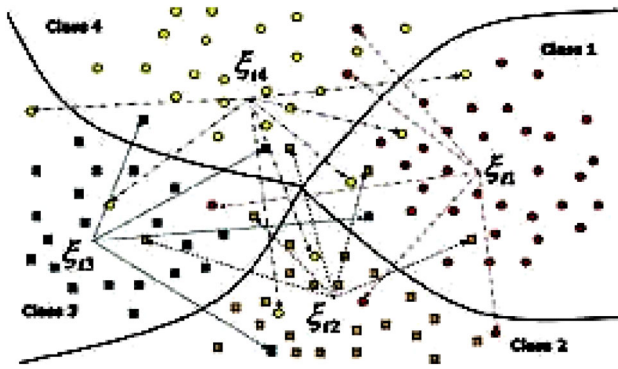


Fig. 5 SVM error representation [35]

imization of

$$Q(w, b, \epsilon) = \frac{1}{2} \sum_{j=1}^k w_j^T w_j + C \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k \epsilon_{ij}, \quad (6)$$

subject to

$$\begin{aligned} w_{y_i}^T \theta(x_i) + b_{y_i} &\geq w_j^T \theta(x_i) + b_j + 1 - \epsilon_{ij} \\ \epsilon_{ij} &\geq 0, i = 1, 2, \dots, n \text{ and } j \in i.2, \dots, k, j \neq y_i, \end{aligned} \quad (7)$$

where x_i is the vector of the i th data point, y_i is the class for the i th data point, ϵ_{ij} is the slack variable as a measure of error, and C is the regularization parameter for balancing error minimization and margin maximization. Figure 5 shows the slack variables for individual classes and formation of the classification problem.

The SVM classifier tries to reduce the value of $k \times n$ slack variables while maximizing the k margins [35]. The multi-class classification function is represented by $\text{argmax}_{j=1, \dots, k} w_j^T \phi(x_i) + b_j$, where a data point (x) is associated with class j . The classification score for the point x is maximized by using weights. The constraints represented in Eqs. 6 and 7 are converted into equivalent unconstrained formulation by Lagrange multipliers α_{ij} and β_{ij}

$$\begin{aligned} Q(w, b, \xi, \alpha, \beta) &= \frac{1}{2} \sum_{j=1}^k w_j^T w_j + C \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k \xi_{ij} \\ &- \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k \alpha_{ij} ((w_{y_i} - w_j)^T \phi(x_i) \\ &+ b_{y_i} + b_j - 1 + \xi_{ij}) \\ &- \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k \beta_{ij} \xi_{ij} = \frac{1}{2} \sum_{j=1}^k w_j^T w_j \\ &- \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k z_{ij} (w_j^T \theta(x_i) + b_j - 1) \\ &- \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k (\alpha_{ij} \\ &+ \beta_{ij} - C) \xi_{ij}, \end{aligned} \quad (8)$$

where

$$z_{ij} = \sum_{j \neq y_i; j=1}^k \alpha_{ij} \text{ for } j = y_i; \quad (9)$$

otherwise,

$$= -\alpha_{ij}, \quad (10)$$

and the conditions for optimality are:

$$\alpha_{ij} ((w_{y_i} - w_j)^T \phi(x_i) + b_{y_i} + b_j - 1 + \xi_{ij}) = 0 \quad (11)$$

for

$$\text{for } j \neq y_i, j = 1, \dots, n \quad (12)$$

$$\beta_{ij} \xi_{ij} = 0 \text{ for } j \neq y_i, j = 1, \dots, k, i = 1, \dots, n, \quad (13)$$

in addition to $Q(w, b, \epsilon, \alpha, \beta)$ being minimized in w, b, ϵ (derivatives equal to zero). The dual formulation is obtained by reducing (3)–(6) using the kernel function $k(x,y)=\phi(x)^T \phi(y)$. The dual formulation is to maximize

$$\begin{aligned} Q(\alpha) &= \sum_{i=1}^n \sum_{j \neq y_i; j=1}^k \alpha_{ij} \\ &- \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^k z_{ij} z_{1j} K(x_i, x_1), \end{aligned} \quad (14)$$

subject to

$$\sum_{i=1}^n z_{ij} = 0 \text{ for } j \neq y_i, j = 1, \dots, n \quad (15)$$

$$0 \leq \alpha_{ij} \leq C \text{ for } j \neq y_i, j = 1, \dots, k, i = 1, \dots, n \quad (16)$$

Finally, the decision function for class j is given by:

$$f_j(x) = \sum_{i=1}^n z_{ij} K(x_i, x) + b_j, \quad (17)$$

and the classification is to assign class j to data point x which satisfies $\text{argmax}_{j=1, \dots, k} f_j(x)$.

The memory requirement and processing time affects the performance of an optimization technique. SVM^{light} is an implementation of the SVM classifier. Its efficiency is enhanced by reducing its training time and suitable selection of kernel parameters.

Abe et al. [37] have improvised by including the bias term. The optimization is represented for n slack formulation by Eq. 18.

$$Q(w, b, \xi) = \frac{1}{2} \sum_{j=1}^k w_j^T w_j + C \sum_{i=1}^n \xi_i \quad (18)$$

subject to

$$\begin{aligned} (w_{y_i}^T - w_j^T) \phi(x_i) + b_{y_i} - b_j &\geq 1 - \xi_i \xi_{ij} \geq 0, 1, \dots, n \\ \text{and } j &\in 1, \dots, k, j \neq y_i \end{aligned} \quad (19)$$

the dual formation is to maximize

$$Q\alpha = \sum_{i=1}^n \xi_i - \frac{1}{2} \sum_{i,l=1}^n \sum_{j=1}^k Z_{ij} Z_{lj} \alpha_i \alpha_l K(x_i, x_l) \quad (20)$$

subject to

$$\sum_{i=1}^n Z_{ij} \alpha_i = 0 \text{ for } j \neq (y_i, j = 1, \dots, k) \quad (21)$$

$$0 \leq (n-1)\alpha_i \leq C \text{ for } j \neq y_i, j = 1, \dots, k, i = 1, \dots, n. \quad (22)$$

The class j decision function is given by

$$f_j(x) = \sum_{i=1}^n Z_{ij} \alpha_j K(x_i, x) + b_j. \quad (23)$$

The regularization parameter (C) limits the value of learned weights as shown in Eq. 23. It performs a balancing act between margin maximization and slack minimization [35].

In this paper, SVM^{light} is used as a classifier for multi-class MI EEG signals [38]. Its polynomial kernel parameters are then varied in a range of values, to attain better performance [39]. Parameter selection plays an important role to obtain accurate classification results [19]. The parameters values can be set, which have a direct effect decision boundary of the classifier [20]. This work focuses on selecting the kernel and then optimizing the values of its parameters to improve the interpretation capability of the decision function. It improves the classification accuracy and overall performance of an MI-based BCI system. The SVM^{light} [40] is an implementation of SVM. It is used in this work, as it is designed for optimization problems.

4.4.1 Parameters selection

Parameter optimization of the selected kernel has a significant effect on the efficiency of the SVM classifier for multi-class EEG signals. In this paper, the polynomial kernel is selected due to its generalization capability. The degree of polynomial kernel characterizes the decision boundary. Eq. 24 represents the achieved decision function.

$$f_j(x) = \sum_{i=1}^n Z_{ij} \alpha_j K(x_i, x) + b_j, \quad (24)$$

where b represents the bias term, x_i is the i th feature vector and n is the number of feature vectors. The vector α_i is the parameter decision boundaries and $K(x_i, x)$ is a kernel function. The polynomial kernel function is represented as $K(x_i, x) = (s a^* b + c)^d$. The kernel makes distinction between multiple classes with significant margin, due to its flexibility [41]. C is a regularization parameter, representing the trade-off between maximization of the margin and the error on the training dataset. The grid search method using cross-validation is executed to assign different values to C from a wide range, to maximize average classification accuracy.

5 Proposed approach

The research work introduces a proposed approach based on SVM in the multi-class EEG signal classification as shown in Fig. 6. In this work, appropriate methods were used for implementation of various stages of the BCI system. The ICA is used for signal preprocessing to remove noise and artifacts from the acquired signals. The FBCSP method, which is a

Fig. 6 Flowchart of the proposed approach

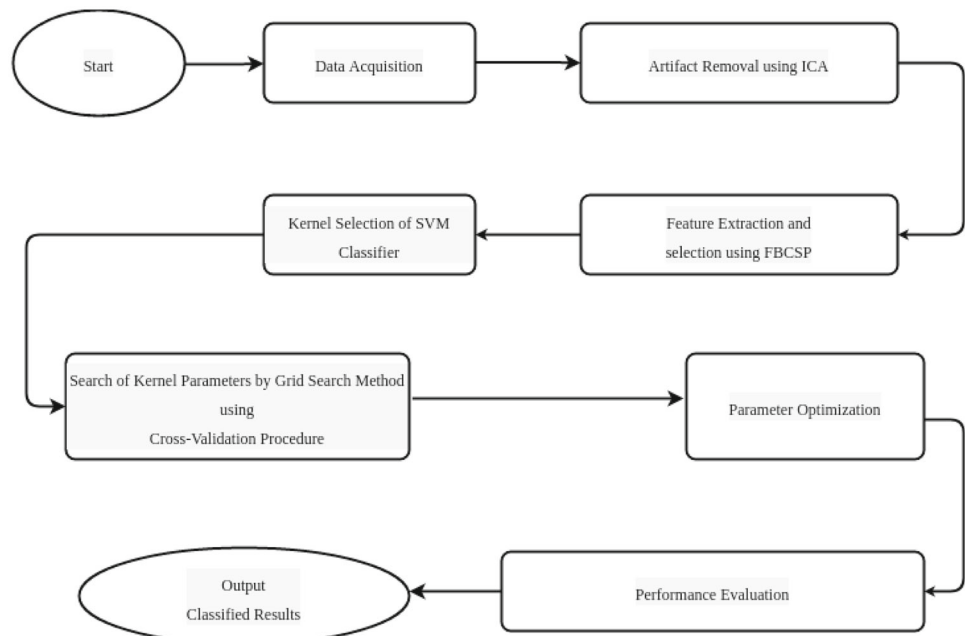


Table 1 Classification accuracy

Subject	$C = [0.1 - 10 - 20 \dots 100]$	$C = [0.01]$	$C = [0.001]$
1	0.722	0.729	0.743
2	0.559	0.563	0.555
3	0.806	0.816	0.823
4	0.417	0.423	0.433
5	0.580	0.569	0.569
6	0.455	0.448	0.473
7	0.875	0.882	0.858
8	0.861	0.886	0.865
9	0.702	0.705	0.719
Average	0.664	0.669	0.670

Fig. 7 Classification accuracy

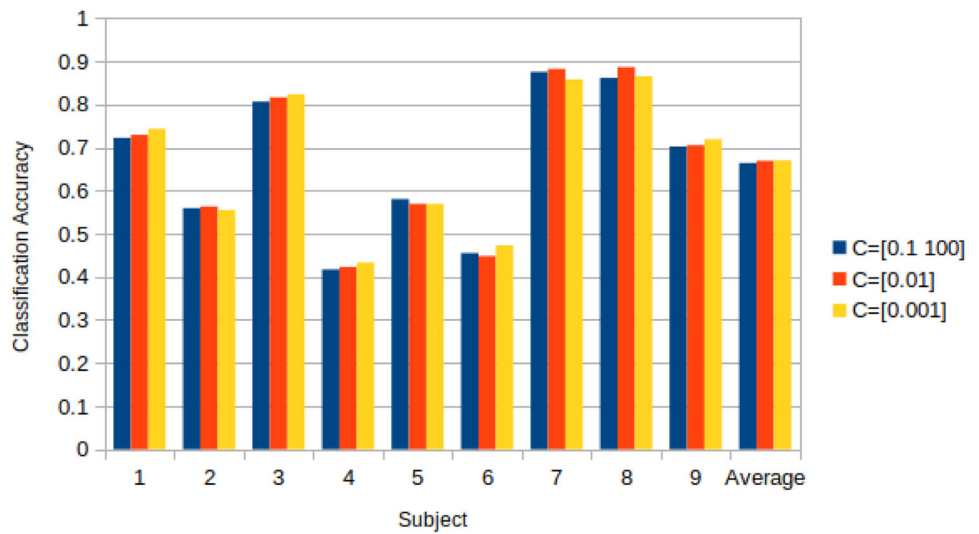


Table 2 Misclassification rate on dataset 2a of BCI competition IV

Subject	$C = [0.1 - 10 - 20 \dots 100]$	$C = [0.01]$	$C = [0.001]$
1	0.278	0.271	0.257
2	0.441	0.437	0.445
3	0.194	0.184	0.177
4	0.583	0.577	0.567
5	0.420	0.431	0.431
6	0.545	0.552	0.527
7	0.125	0.118	0.142
8	0.139	0.114	0.135
9	0.298	0.295	0.281
Average	0.335	0.331	0.329

Table 3 Comparison of related work in literature

Author	Feature Extraction	Classifier	Accuracy%
She et al. (2015)	CSP	SVM	48.4%
		NBPW	53.8%
		NBPW with FBCSP	59.3%
		PPTSVM	62.4%
Proposed approach	FBCSP	SVM-PK	67.0%

Fig. 8 Misclassification rate

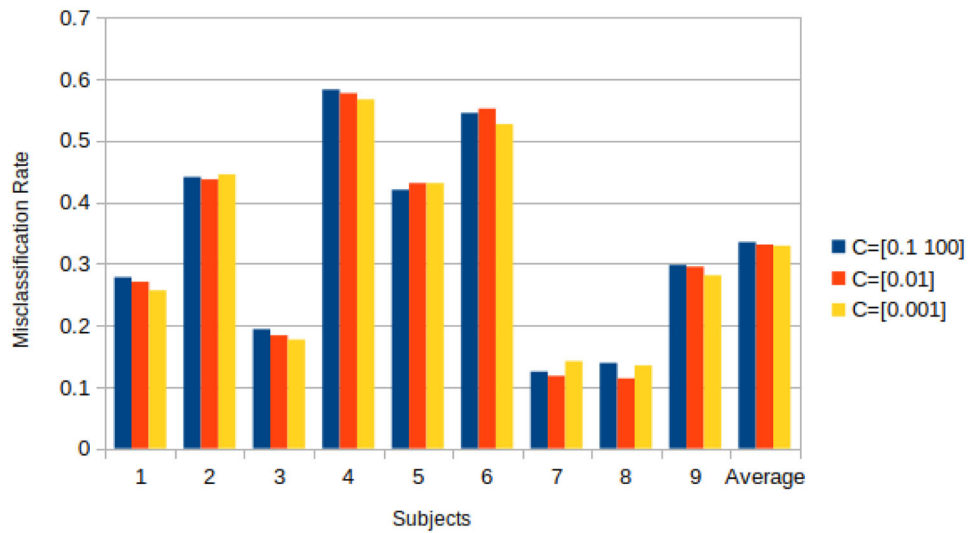
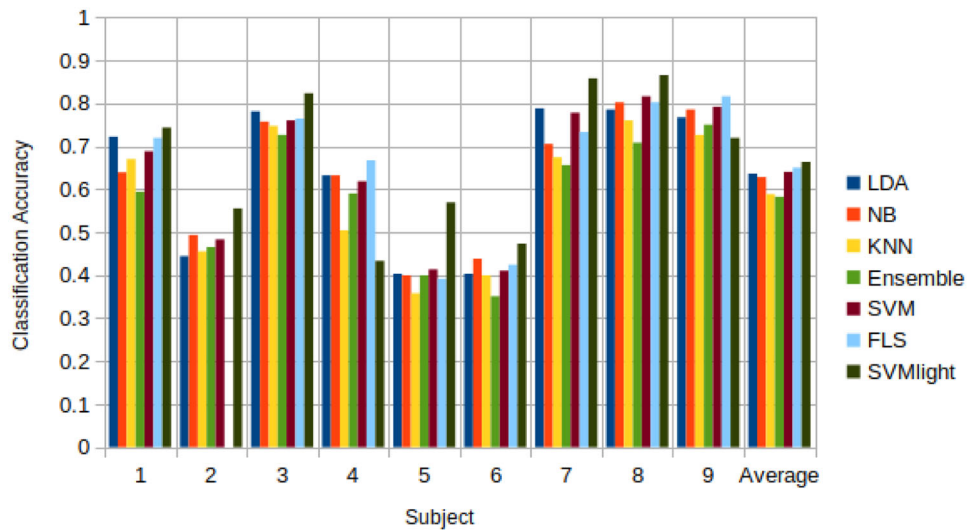


Table 4 Classification accuracy of the proposed approach and existing approaches [21] for BCI IV dataset 2a

Subjects	LDA	NB	KNN	Ensemble	SVM	FLS	SVM-PK
1	0.722	0.639	0.670	0.594	0.688	0.719	0.743
2	0.444	0.493	0.455	0.465	0.483	0.531	0.555
3	0.781	0.757	0.747	0.726	0.760	0.764	0.823
4	0.632	0.632	0.504	0.590	0.618	0.667	0.433
5	0.403	0.399	0.358	0.399	0.413	0.392	0.569
6	0.403	0.438	0.399	0.351	0.410	0.424	0.473
7	0.788	0.705	0.674	0.656	0.778	0.733	0.858
8	0.785	0.802	0.760	0.708	0.816	0.802	0.865
9	0.767	0.785	0.726	0.750	0.792	0.816	0.719
Average	0.636	0.628	0.588	0.582	0.640	0.650	0.670

Fig. 9 Classification accuracy



variant of CSP, is used for extraction and then selection of extracted features. The selected features are then processed by the classifier. Signal classification is performed by using SVM, and its appropriate kernel is chosen. The performance is enhanced by optimizing the parameters of its polynomial kernel. The optimal parameter values were searched using the two-step grid search method. The proposed approach (SVM-PK) is executed on dataset 2a of BCI competition IV, and its performance is evaluated by using the fivefold cross-validation procedure.

6 Results and discussion

The proposed approach is implemented on publicly available dataset 2a of BCI competition IV, and its performance is evaluated. The ICA was used for preprocessing to remove the artifacts. In this paper, FBCSP and SVM-PK are used for feature selection and classification, respectively, for MI-based BCI systems. The performance of the proposed method in terms of different statistical measures, such as classification accuracy and misclassification rate, is evaluated.

In the classifier stage, parameters of the polynomial kernel are assigned to different values to improve the classification accuracy. The degree of polynomial kernel (d) was assigned a value of 3 while varying the regularization parameter (C) by coarse grid search in steps of 10 between 0.1 to 100 [0.1, 10, 20, —, 90, 100]. The classification accuracy of 0.664 was attained at all these values of C . To improve this further, fine grid search was used to find the optimal value of C , by lowering it in the neighborhood of 0.1 and attained a better average classification accuracy of 0.67 for $C = 0.001$ and 0.669 for $C = 0.01$, as shown in table 1. Thus, the classification accuracy was enhanced by finding the optimal values of C and d parameters of the polynomial kernel.

The Fig. 7 shows classification the accuracy for the nine subjects with different C values of the kernel. The misclassification rate was reduced to 0.329 as shown in Table 2 and Fig. 8.

It is shown in Table 3 that the classification accuracy of the proposed approach is improved as compared to its attained value using other approaches as reported in literature.

Classification algorithms of LDA, KNN, NB, Ensemble, FLS and SVM are evaluated, and their performance is reported in literature [21]. Their reported performance on dataset 2a of BCI competition IV is compared with our proposed approach as shown in Table 4 and Fig. 9. It is also compared with classification methods reported in [42], in which the authors have used SVM, NBPW, NBPW with FBCSP and PPTSVM as classifier methods. It is shown that the proposed approach of SVM-PK offers improved classification accuracy for subjects 1, 3, 7, 8 and 9, while the overall average accuracy improved significantly.

7 Conclusion and future scope

In this paper, SVM with polynomial kernel approach is proposed for the classification of multi-class MI EEG signals. The performance of the proposed approach was evaluated from dataset 2a from BCI competition IV. In the preprocessing stage, the ICA is employed for removing artifacts from the acquired EEG signals. The next stage of feature extraction and feature selection is implemented by the FBCSP method. The selected features are then provided to the classifier. Signal classification is performed by SVM with polynomial kernel and its parameters are varied to search for their optimal values, using the grid search method. The performance is evaluated using a fivefold cross-validation process, which indicates the reliability of the obtained results, as it yields better performance. The proposed approach attains an average classification accuracy of 0.67, which is more than that of other approaches executed on the same dataset, as reported in literature. It enhances the reliability and safety of a BCI system for rehabilitation, by improving its classification accuracy. Future research can also investigate different techniques for optimization of different classifier parameters to further improve the efficiency of a BCI system. Future researches can explore other methods for feature extraction and feature selection.

References

1. Dai M, Wang S, Zheng D, Na R, Zhang S (2019) Domain transfer multiple kernel boosting for classification of eeg motor imagery signals. *IEEE Access* 7:49951–49960
2. Zheng Q, Zhu F, Heng P-A (2018) Robust support matrix machine for single trial eeg classification. *IEEE Trans Neural Syst Rehabil Eng* 26(3):551–562
3. Mason SG, Birch GE (2003) A general framework for brain-computer interface design. *IEEE Trans Neural Syst Rehabil Eng* 11(1):70–85
4. Bernardi GB, Pimenta TC, Moreno RL (2019) A simplified tool for testing of feature selection and classification algorithms in motor imagery of right and left hands of eeg signals. In: 2019 IEEE 10th Latin American symposium on circuits & systems (LASCAS). IEEE, pp 197–200
5. Krishna DH, Pasha I, Savithri TS (2016) Classification of eeg motor imagery multi class signals based on cross correlation. *Proc Comput Sci* 85:490–495
6. Berger H (1929) On the eeg in humans. *Arch Psychiatr Nervenkr* 87:527–570
7. Wolpaw J, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan T (2002) Brain-computer interfaces for communication and control. *Clin Neurophysiol* 113(6):767–791
8. Costantini G, Todisco M, Casali D, Carota M, Saggio G, Bianchi L, Abbafati M, Quitadamo L (2009) Svm classification of eeg signals for brain computer interface. In: Proceedings of the 2009 conference on neural Nets WIRN09: proceedings of the 19th Italian workshop on neural nets, Vietri sul Mare, Salerno, Italy, May 28–30 2009. IOS Press, pp 229–233
9. Lotte F, Bougrain L, Cichocki A, Clerc M, Congedo M, Rakotomamonjy A, Yger F (2018) A review of classification algorithms for

- eeg-based brain-computer interfaces: a 10 year update. *J Neural Eng* 15(3):031005
10. Pfurtscheller G, Neuper C, Guger C, Harkam W, Ramoser H, Schlogl A, Obermaier B, Pregenzer M (2000) Current trends in graz brain-computer interface (bci) research. *IEEE Trans Rehabil Eng* 8(2):216–219
 11. Gaur P, Pachori RB, Wang H, Prasad G (2018) A multi-class eeg-based bci classification using multivariate empirical mode decomposition based filtering and Riemannian geometry. *Expert Syst Appl* 95:201–211
 12. Siuly S, Li Y (2012) Improving the separability of motor imagery eeg signals using a cross correlation-based least square support vector machine for brain-computer interface. *IEEE Trans Neural Syst Rehabil Eng* 20(4):526–538
 13. Senthilmurugan M, Latha M, Malmurugan N (2011) Classification in eeg-based brain computer interfaces using inverse model. *Int J Computer Theory Eng* 3(2):274
 14. Padfield N, Zabalza J, Zhao H, Masero V, Ren J (2019) Eeg-based brain-computer interfaces using motor-imagery: techniques and challenges. *Sensors* 19(6):1423
 15. Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B (2007) A review of classification algorithms for eeg-based brain-computer interfaces. *J Neural Eng* 4(2):R1
 16. Mahmood A, Zainab R, Ahmad RB, Saeed M, Kamboh AM (2017) Classification of multi-class motor imagery eeg using four band common spatial pattern. In: 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 1034–1037
 17. Alomari MH, Samaha A, AlKamha K (2013) Automated classification of l/r hand movement eeg signals using advanced feature extraction and machine learning. [arXiv:1312.2877](https://arxiv.org/abs/1312.2877)
 18. Bhuvaneswari P, Kumar JS (2013) Support vector machine technique for EEG signals. *Int J Comput Appl* 63(13):1–5
 19. Yang H, Wu S (2014) Eeg classification for bci based on csp and svm-ga. In: Applied mechanics and materials, vol 459. Trans Tech Publ, pp 228–231
 20. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B (1998) Support vector machines. *IEEE Intell Syst Appl* 13(4):18–28
 21. Nguyen T, Hettiarachchi I, Khatami A, Gordon-Brown L, Lim CP, Nahavandi S (2018) Classification of multi-class bci data by common spatial pattern and fuzzy system. *IEEE Access* 6:27873–27884
 22. Garrett D, Peterson DA, Anderson CW, Thaut MH (2003) Comparison of linear, nonlinear, and feature selection methods for eeg signal classification. *IEEE Trans Neural Syst Rehabil Eng* 11(2):141–144
 23. Kamousi B, Liu Z, He B (2005) Classification of motor imagery tasks for brain-computer interface applications by means of two equivalent dipoles analysis. *IEEE Trans Neural Syst Rehabil Eng* 13(2):166–171
 24. Pfurtscheller G, Muller-Putz G, Schlogl A, Graimann B, Scherer R, Leeb R, Brunner C, Keinrath C, Lee F, Townsend G et al (2006) 15 years of BCI research at Graz university of technology: current projects. *IEEE Trans Neural Syst Rehabil Eng* 14(2):205–210
 25. Ilyas MZ, Saad P, Ahmad MI (2015) A survey of analysis and classification of eeg signals for brain-computer interfaces. In: 2015 2nd international conference on biomedical engineering (ICoBE). IEEE, pp 1–6
 26. Arbabi E, Shamsollahi MB (2017) Evaluation of classical features and classifiers in brain-computer interface tasks. [arXiv:1709.03252](https://arxiv.org/abs/1709.03252)
 27. Zhang D, Yao L, Zhang X, Wang S, Chen W, Boots R, Benatalah B (2018) Cascade and parallel convolutional recurrent neural networks on eeg-based intention recognition for brain computer interface. In: Thirty-second AAAI conference on artificial intelligence
 28. Kumar S, Sharma A, Mamun K, Tsunoda T (2016) A deep learning approach for motor imagery eeg signal classification. In: 2016 3rd Asia-pacific world congress on computer science and engineering (APWC on CSE). IEEE, pp 34–39
 29. Brunner C, Leeb R, Müller-Putz G, Schlogl A, Pfurtscheller G (2008) BCI competition 2008–graz data set a. Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology, vol 16
 30. Tangermann M, Müller K-R, Aertsen A, Birbaumer N, Braun C, Brunner C, Leeb R, Mehring C, Miller KJ, Mueller-Putz G et al (2012) Review of the bci competition iv. *Front Neurosci* 6:55
 31. Liu A, Chen K, Liu Q, Ai Q, Xie Y, Chen A (2017) Feature selection for motor imagery eeg classification based on firefly algorithm and learning automata. *Sensors* 17(11):2576
 32. Ungureanu M, Bigan C, Strungaru R, Lazarescu V (2004) Independent component analysis applied in biomedical signal processing. *Meas Sci Rev* 4(2):18
 33. Bhattacharya A, Bawane DN, Nirxhi S (2015) Brain computer interface using eeg signals. In: GHRCE, Nagpur, vol 17, no 11
 34. Ang KK, Chin ZY, Wang C, Guan C, Zhang H (2012) Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Front Neurosci* 6:39
 35. Bellary J, Eddula KR (2014) Improving multi-class support vector machines training. *Int J Comput Commun Instrum Eng* 1:119–125
 36. Kaur B, Singh D, Roy PP (2018) Eeg based emotion classification mechanism in bci. *Proc Comput Sci* 132:752–758
 37. Abe S (2005) Support vector machines for pattern classification, vol 2. Springer, Berlin
 38. Bandara D, Arata J, Kiguchi K (2018) A noninvasive brain-computer interface approach for predicting motion intention of activities of daily living tasks for an upper-limb wearable robot. *Int J Adv Robot Syst* 15(2):1729881418767310
 39. Ahangi A, Karamnejad M, Mohammadi N, Ebrahimpour R, Bagheri N (2013) Multiple classifier system for eeg signal classification with application to brain-computer interfaces. *Neural Comput Appl* 23(5):1319–1327
 40. Joachims T (1998) Making large-scale svm learning practical. Tech. rep., Technical Report
 41. Ben-Hur A, Ong CS, Sonnenburg S, Schölkopf B, Rätsch G (2008) Support vector machines and kernels for computational biology. *PLoS Comput Biol* 4(10):e1000173
 42. She Q, Ma Y, Meng M, Luo Z (2015) Multiclass posterior probability twin SVM for motor imagery EEG classification. *Comput Intell Neurosci*. <https://doi.org/10.1155/2015/251945>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.