Investigation of EEG Signal Classification Techniques for Brain Computer Interface

Mandeep Kaur Ghumman, Satvir Singh

Abstract: Brain-computer interface (BCI) has emerged as a popular research domain in recent years. The use of electroencephalography (EEG) signals for motor imagery (MI) based BCI has gained widespread attention. The first step in its implementation is to fetch EEG signals from scalp of human subject. The preprocessing of EEG signals is done before applying feature extraction, selection and classification techniques as main steps of signal processing. In preprocessing stage, artifacts are removed from raw brain signals before these are input to next stage of feature extraction. Subsequently classifier algorithms are used to classify selected features into intended MI tasks. The major challenge in a BCI systems is to improve classification accuracy of a BCI system. In

this paper, an approach based on Support Vector Machine (SVM), is proposed for signal classification to improve accuracy of the BCI system. The parameters of kernel are varied to attain improvement in classification accuracy. Independent component analysis (ICA) technique is used for preprocessing and filter bank common spatial pattern (FBCSP) for feature extraction and selection. The proposed approach is evaluated on data set 2a of BCI Competition IV by using 5-fold crossvalidation procedure. Results show that it performs better in terms of classification accuracy, as compared to other methods reported in literature.

Keywords: Brain computer interface, electroencephalography, motor imagery, filter bank common spatial pattern, support vector machine, independent component analysis.

I. INTRODUCTION

The brain-computer interface (BCI) is a method of establishing a communication channel between a user and system, which is independent of brain's normal output nerve pathways and muscles [1]. It provides an advanced technology which can translate intent of a user from brain signals directly into commands and can thus establish a direct communication channel between human brain and external devices [2]. The most used signal acquisition technique in BCI studies is the electroencephalography (EEG) due to its simplicity and usability [3]. EEG provides high temporal

resolution at low cost, making it popular among researchers [4]. In EEG based BCI systems, noninvasive sensors are placed on the scalp of user to sense electrical activity of the brain [5]. The EEG signals are categorized into subcategories of EEG sub-bands - delta, theta, alpha, beta and gamma based on frequency of signals [6]. The standardized international 10-20 electrode placement system of EEG demarcates the position of electrodes on various parts of the subject's scalp [7].

   Mandeep Kaur Ghumman, Research Scholar, IKG Punjab Technical University Kapurthala, Punjab, India Email: mghumman20@yahoo.com
   Satvir Singh, Associate Professor, IKG Punjab Technical University Kapurthala, Punjab, India Email: drsatvir:in@gmail.com

It is used to record signals emanating as byproduct of a mental activity, from somatosensory and the motor areas of brain. The signals recorded from these electrodes, reflect the motor activity of the person like that of hand, foot and tongue movements etc. [8]. Applications of BCI are found in areas of robotics, mobility devices, environmental control, and device communication. It has the capability to provide an alternative communication as well as control technology for rehabilitation, to patients suffering from severe neuromuscular disorders [9]. A popular paradigm for BCI communication is motor imagery (MI), which is a common mental task wherein a subject is instructed to imagine movement of a particular limb. In such a system, EEG signals are recorded during execution of multiple MI tasks (e.g. right hand, left hand, foot and tongue movement etc.) [10]. Various aspects of MI based BCI have been intensively studied, for which some benchmark data sets are used to measure performance of various approaches in BCI research. BCI system is composed of sequential stages of signal acquisition, signal preprocessing, feature extraction and classification [11]. In stage of signal acquisition, the electrical activity created by brain is recorded from scalp of the subject and is read using electrodes while subject performs specified voluntary task [12]. However, noise and artifacts can interfere and thus influence the recorded EEG data. These are undesired signals which can introduce undesirable changes in the fetched brain signals [13]. These are induced by power line noise, changes in electrode impedance and potentials introduced by movements of body parts. In this paper, independent component analysis (ICA) [14] is used as a preprocessing method to remove artifacts. The architect of a BCI system faces a big challenge to use appropriate algorithms for its different stages, to efficiently identify EEG signals corresponding to different MI tasks [15]. Subsequently feature extraction method is used to isolate

and extract distinguishing features from the brain signals. Common spatial pattern (CSP) is a powerful motor imagery feature extraction method used for classification problems [16]. The high dimensionality of extracted feature vector adversely affects the classifier performance [17]. Hence feature selection is used to select only relevant features [18]. In this paper, optimization of spatial and spectral filters is done by filter bank common spatial pattern (FBCSP), which is an extension of CSP [19]. It makes use of several filter banks to select a reduced set of features from predefined narrow bands. Then CSP algorithm is used to harvest spatial filters from each band-pass filtered EEG signals. Feature selection is applied using Mutual Informationbased Best Individual Feature (MIBIF) to obtain suitable subset of features [20].

These features are classified to a particular class by the classifier.

The architect of a BCI system has to finally make choice of a suitable classifier from some of well-known classifiers like support vector machine (SVM), linear discriminant analysis (LDA), k-nearest neighbor algorithm (KNN), fuzzy logic (FL) and artificial neural network (ANN) for EEG classification [21]. For multi-class classification, selection of a particular classifier is a critical issue in BCI system [22]. Linear classifiers are generally preferred for EEG classification due to their low computational complexity and better stability [23].

They are also less prone to over fitting problem as compared with non-linear classifiers, especially when only a limited number of samples are available [24]. SVM is one of the most popularly applied classifiers for BCI systems. It works by establishing an optimum hyperplane separating different classes to a maximum possible extent [25]. It is able to execute multi-class classification and is independent of the dimensional of data. Selection of optimal kernel parameters is of paramount importance to obtain accurate classification results [26]. Varying the parameters setting effects the boundary decision in the classifier [27]. This paper presents selection of suitable kernel and optimization of parameters of kernel to enhance the interpretability of the decision function. It improves the classification accuracy and overall performance of BCI system.

II. MAJOR CONTRIBUTIONS

In this paper, SVM approach is used for signal classification of data set 2a of BCI Competition IV. Parameters of kernel are found by means of grid-search method and are varied to enhance the classification accuracy. K-fold cross validation was used to evaluate performance of the classifier. This study has improved the classification accuracy of MI data in a BCI system, which is verified and compared with other methods as reported in the literature [28].

III. RELATED WORK

Many modern algorithms can be used to implement various stages of a BCI. Many researches have reported a variety of methods using datasets of BCI Competition III and IV. Lotte et al. have reviewed various classification algorithms in BCI context to identify their critical properties [23]. They have analyzed and compared performance of linear classifiers, NNs, nonlinear Bayesian classifiers, nearest neighbor classifiers and combinations of classifiers. They have provided guidelines for choosing a classification algorithms for specific BCIs. They have observed that SVM are more suitable for synchronous BCIs due to its regularization property and immunity to curse-of-dimensionality. The authors in [29] have reported a clustering techniquebased least square SVM algorithm (LS-SVM) to perform EEG signal classification. They have developed a clustering technique to implement feature extraction and then used the obtained features as input to LS-SVM for classification. They have improved classification accuracy and evaluated its performance using 10-fold cross-validation method. The limitation of their approach was manual selection of parameters for LS-SVM method. Selim et al. have

introduced bio-inspired algorithms based approach for feature selection and classifier optimization to increase classification accuracy of the MI-BCI systems [30]. They have selected optimized time interval for each subject and extracted features from EEG signals using CSP. They have used a hybrid attractor metagene (AM) algorithm and Bat optimization algorithm (BA) for selection of most discriminant CSP features and optimization of SVM parameters. Their approach outperformed on benchmark multi-class data sets. Higher time taken by execution of BA is a major drawback of their proposed approach. They have suggested optimization to reduce training time in future research. They have also suggested use of deep learning approaches besides evaluation on new data sets. The authors in [28] have proposed a fuzzy logic system (FLS) based approach for multi-class MI data classification. They have fused fuzzy system with particle swarm optimization (PSO) method for improving classification performance. They have used CSP algorithm in feature extraction phase to extract relevant discriminant features from multi-class EEG data. The learning process of an FLS is computationally intensive. Hence they have reduced the computational expense of the multi-class FLS based BCI system by application of PSO to reduce processing time. They have cross-validated performance of the proposed FLS method on benchmark data sets. They have suggested to study more efficient feature extraction and selection methods in future research to improve the classification performance of a BCI system. A method which uses CSP for feature extraction and SVM with genetic algorithm (SVM-GA) for classification, has been proposed by authors in [26]. They have used GA to optimize kernel parameters setting. The proposed method was evaluated on dataset IVa of BCI Competition III, wherein it outperformed the conventional LDA in classification performance. They have recommended study of different feature selection methods in future research. The authors in their research [24] have utilized frequency ranges of mu and beta rhythms of EEG signals . They have used CSP for feature extraction and SVM for classification. The have implemented their approach on dataset IIIa of BCI competition III, and shown improvement in classification accuracy. They have suggested use of proposed approach for implementation of online BCI system. Siuly et al. reported a approach for classification of MIbased EEG signals in BCIs [15]. They have used approach of cross-correlation feature extraction and least square support vector machine (LS-SVM) for classification of twoclass MI EEG signals. The classification accuracy of their proposed method was evaluated on datasets IVa and IVb of BCI competition III, through a 10-fold cross-validation procedure. They have used a twin step grid search algorithm for identifying optimal combinations of parameters for LSSVM classifier. They have suggested to extend this approach for multi-class problems in future research.

IV. MATERIALS AND METHODS

4.1. Dataset

In this paper dataset 2a from BCI Competition IV is used, which is publicly available for researchers [31]. The dataset has recordings of EEG signals from nine subjects while they performed MI tasks. The data acquisition consisted of 22 EEG besides 3 EOG channels as shown in figure 1. The electrodes arranged as per standard 10-20 system across scalp of the human subject.

The data is then stored in general data format (GDF) for biomedical signals, consisting of one file for each of the subjects. The subjects were seated in chairs, facing a computer screen, and signaled to perform a desired MI task movement of left hand, right hand, both feet or tongue. The data was collected for two sessions. Each session consisted of 6 runs punctuated by short breaks, with each run consisting of 48 trials of 12 for each of the four possible MI classes. It generated a total of 288 trials per session, with 72 trials corresponding to each of the 4 classes.

The timing of data acquisition is shown in figure 2, wherein the beginning of each trial is commenced by appearance of a fixation cross on the black screen, along-with a short acoustic warning tone. Then a symbol of a left, right, up or down arrow is displayed on the screen, corresponding to one of the four classes of left hand, right hand, foot and tongue MI task. This is a signal to the subject to perform the desired MI task. The signals are sampled at 250 Hz and then band-pass filtered between 0.5 Hz and 100 Hz. In addition to this, a 50 Hz notch filter is used to suppress noise from power line. Here EOG channels are provided for the subsequent application of artifact processing methods [32]. The signals then are preprocessed to increase signal-tonoise ratio and remove artifacts, before application of feature extraction method.

4.2. Preprocessing

Acquisition of data is prone to noise from external environment, so preprocessing is required to remove these artifacts from EEG signals. The efficiency of preprocessing Figure 1. Electrodes Location by International 10-20 System [20] Figure 2. Timing scheme of the paradigm for data set 2a from BCI Competition IV [33] phase has a direct impact on efficiency attainment of overall BCI system [34]. The raw data set stored in GDF format was loaded by using functions of BioSig toolbox [35]. Artifacts were removed from EEG signals using EEGLAB [36] [37], which is an interactive MATLAB toolbox used for preprocessing of EEG signals. ICA was employed for artifact removal from acquired data to obtain relevant information contained in signals. ICA is a computation method for separation of multi-sourced signals into appropriate sub components, with an assumption that signals are statistically independent [13]. It is based on blind source separation of independent components by increasing the statistical independence of estimated components to a maximum value. ICA has capability of revealing information from higher order statistics of data. Technique of ICA is usually applied for removal of EOG, EMG and ECG artifacts. In our dataset, 22 EEG channels besides 3 EOG channels were used to record data. ICA is employed to remove 3 EOG channels

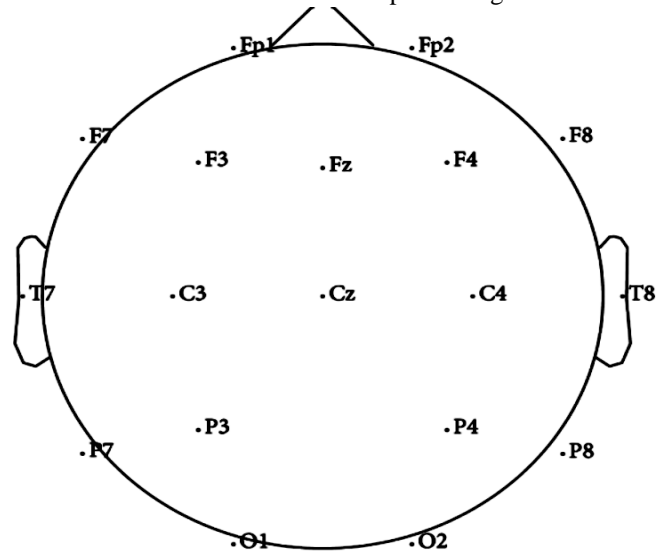corresponding to eye movement [38], and remaining 22 EEG channels were used for further processing.

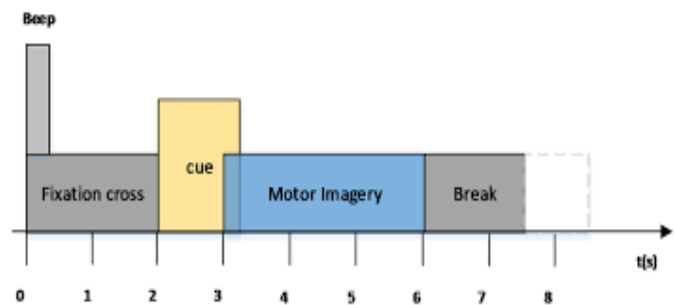**Figure 1. Electrodes Location by International 10-20 System [20]**

Figure 2. Timing scheme of the paradigm for data set 2a from BCI Competition IV [33]

phase has a direct impact on efficiency attainment of overall BCI system [34]. The raw data set stored in GDF format was loaded by using functions of BioSig toolbox [35]. Artifacts were removed from EEG signals using EEGLAB [36] [37], which is an interactive MATLAB toolbox used for preprocessing of EEG signals. ICA was employed for artifact removal from acquired data to obtain relevant information contained in signals. ICA is a computation method for separation of multi-sourced signals into appropriate sub components, with an assumption that signals are statistically independent [13]. It is based on blind source separation of independent components by increasing the statistical independence of estimated components to a maximum value. ICA has capability of revealing information from higher order statistics of data. Technique of ICA is usually applied for removal of EOG, EMG and ECG artifacts. In our dataset, 22 EEG channels besides 3 EOG channels were used to record data. ICA is employed to remove 3 EOG channels corresponding to eye movement [38], and remaining 22 EEG channels were used for further processing.
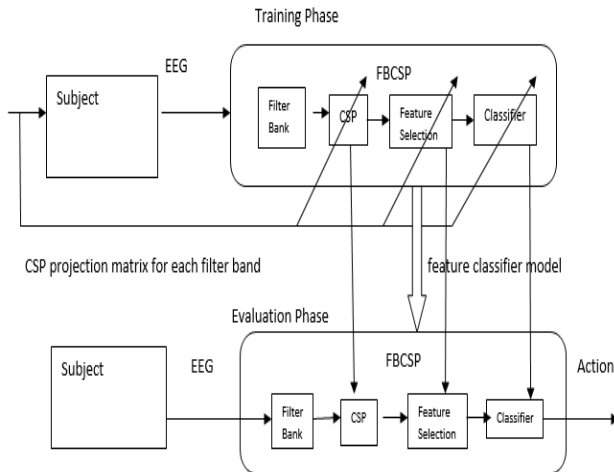
Figure 3. Filter Bank Common Spatial Pattern [39]

4.3. Feature Extraction

After preprocessing, feature extraction process retrieves the most relevant features from input signal. It generates more precise description of these features making it suitable for further processing. These features are required for decision making mechanism in generating the required output. Spatial filters and temporal/spectral filters are used for distinguishing patterns of different MI signals. Optimization of spatial and spectral filters plays an important role in improving the performance of a BCI system. FBCSP [39] was used as a feature extraction method. The FBCSP algorithm consist of signal processing and execution of machine learning procedure on EEG data.

It consist of four progressive stages - filter bank consisting of band pass filters, spatial filtering by use of CSP method, MIBIF for feature selection and classification of the selected relevant features, as shown in figure 3. Band-pass filtering is the preliminary phase of FBCSP which makes use of filter bank for decomposing EEG signals into various frequency pass bands by use of the causal Chebyshev Type II filter. It uses nine band pass filters, which falls in range of 4- 8, 8- 12,..... and 36-40 Hz, to achieve a stable frequency response. Spatial filtering is a second phase of FBCSP and uses the CSP algorithm as it offers higher efficiency in calculation of spatial filters. These are further used for detection of event-related desychronization/synchronization (ERD/ERS). The band pass and spatial filter implements the spatial filtering of EEG signals, which have previously been filtered with a particular frequency range. Each of such pair of the band pass filter and spatial filter, calculates CSP features belonging to a particular band pass frequency range. The stage of spatial filtering is executed by the CSP method, which linearly transforms EEG signal using

$$Z_{b;i} = W_{T\,b}\, E_{b;i} \quad (1)$$

where $E_{b;i}\, 2\, R_{cxt}$ represents single trial EEG signal from b_{th} band pass filter from i_{th} trial; $Z_{b;i}\, 2\, R_{cxt}$ denotes $E_{b;i}$ after spatial filtering. $W_b\, 2\, R_{cxc}$ represents the CSP projection matrix, c denotes number of channels, t denotes number of EEG samples per channel, and T denotes the transpose operator.

The CSP algorithm calculates transformation matrix Wb, to obtain features with variances which are optimal for

distinguishing two classes of the EEG signals, by resolving the eigen value decomposition problem

$$\Sigma_{b,1}\, W_b = (\Sigma_{b,1} + \Sigma_{b,2})\, W_b D_b\, (2)$$

in which, $\Sigma_{b,1}$ and $\Sigma_{b,2}$ represents the covariance matrices of EEG signals of the respective MI action which were filtered by b^{th} band pass, and D_b is the diagonal matrix containing the eigen values of $\Sigma_{b,1}$. The value of W_b is calculated in MATLAB by executing the command W = eig(S1, S1+S2), where W represents W_b, S1 represents $\Sigma_{b,1}$ and S2 represents $\Sigma_{b,2}$. The value of difference between variance of the two class EEG signals, which is band pass filtered, is maximized by use of W_b from equation 2. Thus, the m-pairs of CSP features from i^{th} trial of bth band pass filtered EEG signals is represented by

$$v_{b,i} = log\frac{diag(\overline{W_b}^T E_{b,i} E_{b,i}^T \overline{W_b})}{tr[\overline{W_b}^T E_{b,i} E_{b,i}^T \overline{W_b}]} \quad (3)$$

where $v_{b,i}\, R^{2m}$, W_b denotes first m and last m columns of W_b, diag(.) represents diagonal elements of square matrix, tr[.] represents the summation of diagonal elements of square matrix. and m is assigned value of 3 for dataset 2a used in the experiment. The FBCSP feature vector, is then calculated for the ith trial in the experiment as following

$$v_i = [v_{1,i}, v_{2,i}, \ldots\ldots, v_{9,i}];\,(4)$$

where

$$v_i \in R^{1\,X(9*2m)}, i = 1,2,3,\ldots\ldots, n;\,(5)$$

where n represents number of trials executed in data under consideration and $V \in R^{n_t X(9*2m)}$; $y \in R^{n_t X1}$; v_i and y_i represents feature vector and actual class label from the ith training session , i = 1,2............n_t ; and n_t represents the total number of trials executed in training data set under consideration. Analysis of EEG generates a huge quantity of data. The EEG signals are recorded with sampling frequency of 250 Hz. These signals have to be converted into a fewer number of values which characterize some particular features of the signals e.g. power of the EEG signals, in various frequency bands. Such features are further processed and aggregated into a feature vector. The selection and extraction of relevant features improves the performance of the system by easing task of classification algorithm. Feature selection is executed on the training data, by selection of distinguishing CSP features based on relation

MIBIF method is based on the filter approach.

The mutual information of each feature is calculated and then sorted in decreasing order. The first k features are then selected, and are subsequently used for classification stage.

4.4. Classification

The classification stage of BCI executes a classification algorithm to model the selected features and classify it into
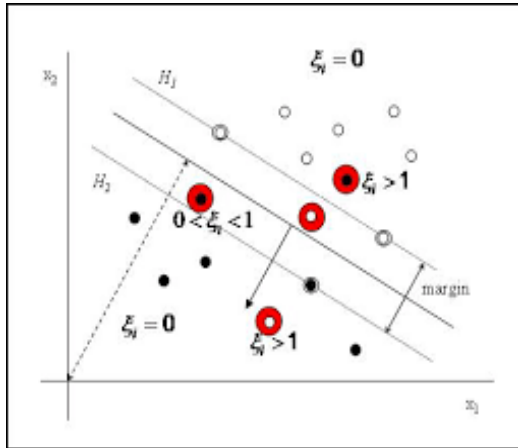

Figure 4. SVM Non-Linearly Separable data points [42]

corresponding MI classes. It performs automatic association of a MI class to an appropriate feature vector, which is extracted in the previous stage. This class identifies the type of MI task performed by the BCI user. SVM is considered as a suitable classifier for BCI systems to perform classification of MI-EEG signals. SVM classifies data by establishing a hyperplane which separates all data points of one class from those of the other classes. SVM establishes decision boundaries with help of support vectors and separates multiple classes by mapping of data to a higher dimensional space with maximum margins by means of a kernel function [40]. SVM is flexible and allows adjustment of many parameters to improve classification rate. It is also preferred for its ease of use, and for offering higher success rate for classification. The SVMlight [41] is a special purpose solver for SVM optimization problems. In this paper, SVM light is used as a classifier.

4.5. Support Vector Machine

SVMlight has been successfully employed in many MI based BCIs. It has a fast optimization algorithm based on selection of working set and caching of kernel evaluations. It uses sparse vector representation, and thus can handle thousands of support vectors and training examples. It also supports standard kernel functions besides creation of customized kernel functions [43]. The SVM performs classification by establishing linear decision hyperplanes in feature space for separating the training data using different labels during training state. The training data is separated using a kernel function when it is otherwise not possible to separate it in a linear manner. In this paper, SVM is used for classification of multi-class (MI) EEG signals [44]. In proposed approach, different parameters can be assigned appropriate values for improving its performance [45]. The values of regularization parameter (C), gamma ( ) and

degree of kernel (d) are selected to control the trade-off between number of non-separable points and complexity of algorithm.
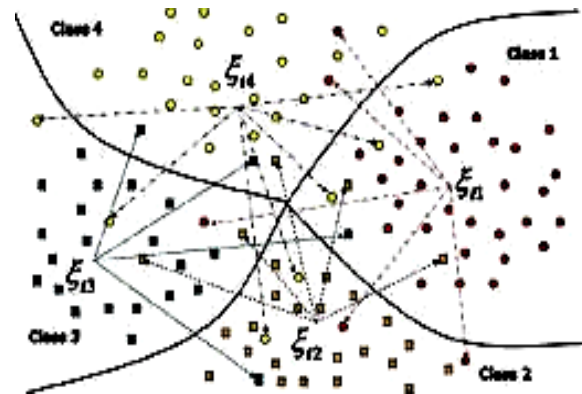

Figure 5. Multi-Class SVM error representation [42]

In case of non-separable data, SVM attempts to maximize the margin separating various classes while reducing classification error to a minimum value for each data point, represented by the slack variable, as shown in figure 4. For a k-class problem with n training points, the SVM can be expressed as a minimization of

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

subject to

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

where x_i is the input vector for data point i, y_i is the appropriate class of data point i, s_{ij} is slack variable of data point i related to class j representing measure of error, C is regularization parameter denoting trade-off between margin maximization and error minimization in classification. Fig- ure 5 shows different slack variables specific to individual classes and formulation of multi-class classification prob- lem.

In multi-class SVM formulation, the classifier attempts to minimize value of k x n slack variables while attempt- ing to maximize the k margins [42]. The multi-class clas- sification decision function is defined by argmax$_{j=1\ \ k}$

w$^T_i\varphi(x_i) + b_j$ where a data point x is classified to class

j with weights maximizing the classification score for the point x. The problem of constraints represented in (1) and (2) can be converted into its unconstrained equivalent formu- lation by using non-negative Lagrange multipliers a_{ij} and β_{ij}

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

where

$$z_{ij} = \Sigma_{?,...,}^{k} \;.\alpha_{ii} \; for \; i = u_i \tag{9}$$

otherwise

$$= -\alpha_{ij} \tag{10}$$

and the conditions for optimality are:

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

for

$$for \, j \neq y_i, j = 1...n \tag{12}$$

$$\beta_{ij}\xi_{ij} = 0 \; for \; j \neq y_i, j=1....k, i=1......n \tag{13}$$

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) to (6) using the kernel function $k(x,y)=\phi(x)^T\phi(y)$. The dual formulation is to maximize

$$Q(\alpha) = \Sigma_{i=1}^{n}\Sigma_{j\neq y_i;j=1}^{k}\alpha_{ij} - \frac{1}{2}\Sigma_{i=1}^{n}\Sigma_{j=1}^{k}z_{ij}z_{1j}K(x_i,x_1) \tag{14}$$

subject to

$$\Sigma_{i=1}^{n}z_{ij} = 0 \; for \; j \neq y_i, j=1......n \tag{15}$$

$$0 \leq \alpha_{ij} \leq C \; for \; j \neq y_i, j=1.....k, i=1....n \tag{16}$$

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

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

and the classification task for data point x is to find class j to satisfy $argmax_{j=1:::k}f_j(x)$. The efficiency of an optimization technique depends upon many factors such as processing time and memory requirements. SVM$_{light}$is an implementation of SVM learner and its computational efficiency can be improved by reduction in training time and appropriate selection of kernel parameters.

4.5.1. Improving SVM Training Time. In this paper, an proach proposed by Crammer et al. [46] is used to reduce training time by optimizing number of variables and/or constraints. The size of the optimization problem is reduced by reducing the number of slack variables and selecting highest slack for each data point across all classes. Abe et al. [47] have extended this formulation to include bias term. Given the n slack formulation, optimization problem is represented by

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

subject to

$$(w_{yi}^T - w_j^T)\phi(x_i)+b_{yi}-b_j \geq 1 - \xi_{ij}\xi_{ij} \geq 0, 1,.....n$$
$$and \, j \in 1,.....k, j \neq y_i \tag{19}$$

the dual formation is to maximize

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

subject to

$$\Sigma_{i=1}^{n}Z_{ij}\alpha_i = 0 \; for \; j \neq (y_i, j=1,....k) \tag{21}$$

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

the class j decision function is given by

$$f_{j(x)} = \Sigma_{i=1}^{n}Z_{ij}\alpha_j K(x_i,x)+b_j \tag{23}$$

It is observed that a reducing slack variables from n x k to n does not lead to increase in number of constraints. The final optimization problem is also simplified as compared to basic SVM problem. All learned weights are limited by the value of the regularization parameter C as given by learning constraint in equation 24. It implements a trade-off between slack minimization and margin maximization [42].

4.5.2. Kernel Parameter Settings. The performance of the SVM for classification of four-class EEG signals depends on the choice of a kernel. Optimization of parameters of a kernel can train our classifier for a given dataset and improve classification accuracy of a classifier [48]. In this paper, appropriate kernel is chosen after study and performing many trials. The degree of Polynomial Kernel has a direct influence on flexibility of resulting classifier. The Polynomial Kernel is a global kernel which has a good generalization ability. Parameters of a kernel have a significant effect on the decision boundary. It can classify data with nonlinear boundaries as well as of high dimensions. A final decision function is achieved using the following equation

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

where x$_i$ denotes ith input feature vector of d dimensions, n is the number of feature vectors and b represents bias term. The vector a_i includes the parameters defining decision boundaries in the kernel space and K(x$_i$,x) represents a kernel function.The RBF kernel function is defined as K(x$_i$, x) = $(\; gamma \; a \; b \;^2)$. The kernel is flexible enough to discriminate between two or multiple classes with a sizable margin [49]. The Polynomial Kernel function is defined as K(x$_i$, x) = $(s\, a{*}b + c)^d$, where d is the degree of kernel. The regularization parameter, denoted by C, is a trade-off between error on training data set and margin maximization. The value of C is randomly varied through a wide range using grid search method with cross-validation to attain maximum average classification accuracy.

V. RESULTS AND DISCUSSION

This paper proposes a FBCSP/SVM machine learning approach for multi-class MI based BCI systems. The ublicly accessible dataset 2a from BCI Competition IV was used, which was recorded from nine subjects while they performed MI tasks. The data was collected for two essions. It generates a total of 288 trials per session, with 72 trials corresponding to each of the 4 classes. The fetched EEG signals were preprocessed using ICA to remove the artifacts. The feature extraction and selection was done using FBCSP, while SVM was used as a classifier technique. The performance of proposed
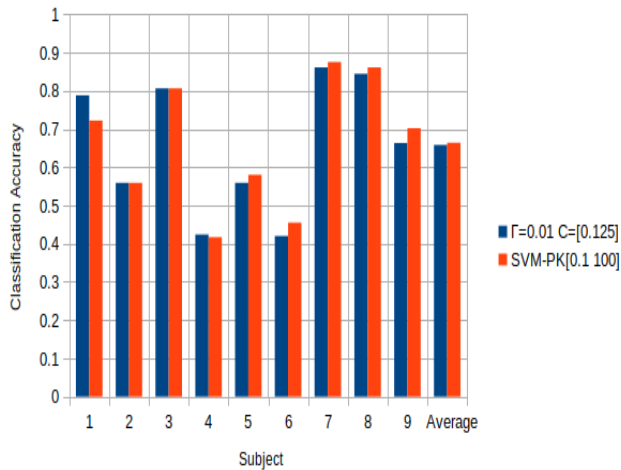
approach was calculated in terms of classification accuracy.

Table 1. Classification Accuracy On Dataset 2a Of Competition Iv

Subject	Radial Basis Function Kernel at C=[0.125] γ=0.01	Polynomial Kernel at C= [0.1 100] d=3
1	0.788	0.722
2	0.559	0.559
3	0.806	0.806
4	0.424	0.417
5	0.559	0.580
6	0.42	0.455
7	0.861	0.875
8	0.844	0.861
9	0.663	0.702
Average	0.658	0.664

Figure 6. Classification accuracy

In classifier stage, SVM was used and parameters of the Radial Basis Function (C and γ) and Polynomial (d) kernels were varied to improve the classification accuracy. SVM-RBF attained best classification accuracy at C=0.125 and γ=0.01. For Polynomial Kernel, the value of regularization parameter (C) was randomly varied between 0.1 and 100 resulting in attainment of maximum average classification accuracy of 0.664 at C=0.1 and d=3, as shown in table 1.

The figure 6 depicts classification accuracy results for each of the 9 subjects. The whole data set is cross-validated at optimal value of C.

The classification accuracy of the proposed approach was found to be better as compared to other approaches reported in literature, and is shown in table 2.

We have compared the proposed approach with other approaches namely LDA, NB, KNN, ensemble AdaBoostM2, SVM and PSO based FLS as reported in literature [28]. The performance of each classifier in terms of the maximum accuracy rates, executed on test dataset of each subject, is shown in table 3 and figure 7. It can be observed that our proposed approach has led to increase in average classifica-

Table 2. "Comparison of Related Work In Literature"

Author	Feature Extraction	Classifier	Accuracy%
She et al. (2015)	CSP	SVM	48.4%
		NBPW	53.8%
		NBPW with FBCP	59.3%
		PPTSVM	62.4%
Proposed approach	FBCSP	SVM-PK	66.4%

Table 3. "Classification Accuracy Of The Proposed Approach And Existing Approaches For Bcic 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	0.722
2	0.444	0.493	0.455	0.465	0.483	0.531	0.559
3	0.781	0.757	0.747	0.726	0.760	0.764	0.806
4	0.632	0.632	0.504	0.590	0.618	0.667	0.417
5	0.403	0.399	0.358	0.399	0.413	0.392	0.580
6	0.403	0.438	0.399	0.351	0.410	0.424	0.455
7	0.788	0.705	0.674	0.656	0.778	0.733	0.875
8	0.785	0.802	0.760	0.708	0.816	0.802	0.861
9	0.767	0.785	0.726	0.750	0.792	0.816	0.702
Average	0.636	0.628	0.588	0.582	0.640	0.650	0.664
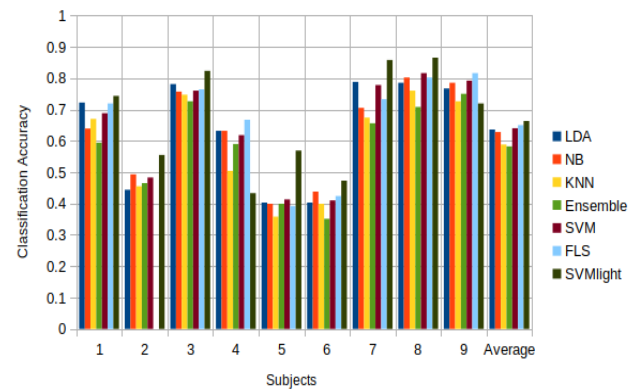
Figure 7. Classification accuracy of the proposed approach and existing approaches

tion accuracy. This paper presents better results as compared to other approaches used in previous literature.

VI. CONCLUSION AND FUTURE SCOPE

In this paper, the SVM is used for classification of multiclass MI EEG signals. In preprocessing, the ICA is used to remove artifacts from EEG signals and FBCSP method is used for feature extraction and selection. The selected features are input to the classifier, where signal classification is done using SVM while optimizing kernel parameters of RBF and PK, to improve classification accuracy. The performance of proposed approach is evaluated on dataset 2a of BCI Competition IV, using 5-fold cross-validation procedure. The paper has shown an improvement in average classification accuracy as compared to other approaches reported in literature. Different parameter optimization and feature selection techniques will be investigated in future research.

REFERENCES

1. M. Dai, S. Wang, D. Zheng, R. Na, and S. Zhang, "Domain transfer multiple kernel boosting for classification of eeg motor imagery signals," IEEE Access, vol. 7, pp. 49951–49960, 2019.
2. Q. Zheng, F. Zhu, and P.-A. Heng, "Robust support matrix machine for single trial eeg classification," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, no. 3, pp. 551–562, 2018.
3. G. B. Bernardi, T. C. Pimenta, and R. L. Moreno, "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), pp. 197–200, IEEE, 2019.
4. D. H. Krishna, I. Pasha, and T. S. Savithri, "Classification of eeg motor imagery multi class signals based on cross correlation," Procedia Computer Science, vol. 85, pp. 490–495, 2016.
5. S. Kumar, A. Sharma, K. Mamun, and T. Tsunoda, "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), pp. 34–39, IEEE, 2016.
6. I. Dokare and N. Kant, "Classification of eeg signal for imagined left and right hand movement for brain computer interface applications," International Journal of Application or Innovation in Engineering & Management, vol. 2014, pp. 291–294, 2014.
7. G. Costantini, M. Todisco, D. Casali, M. Carota, G. Saggio,
8. L. Bianchi, M. Abbafati, and L. Quitadamo, "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, pp. 229–233, IOS Press, 2009.
9. F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rako-tomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," Journal of neural engineering, vol. 15, no. 3, p. 031005, 2018.
10. J. C. Lee and D. S. Tan, "Using a low-cost electroencephalograph for task classification in hci research," in Proceedings of the 19th annual ACM symposium on User interface software and technology, pp. 81–90, ACM, 2006.
11. P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "A multi-class eeg-based bci classification using multivariate empirical mode decompo-sition based filtering and riemannian geometry," Expert Systems with Applications, vol. 95, pp. 201–211, 2018.
12. M. R. Lakshmi, T. Prasad, and D. V. C. Prakash, "Survey on eeg signal processing methods," International Journal of Advanced Re-search in Computer Science and Software Engineering, vol. 4, no. 1, 2014.
13. J. Thomas, T. Maszczyk, N. Sinha, T. Kluge, and J. Dauwels, "Deep learning-based classification for brain-computer interfaces," in 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 234–239, IEEE, 2017.
14. M. Ungureanu, C. Bigan, R. Strungaru, and V. Lazarescu, "Indepen-dent component analysis applied in biomedical signal processing," Measurement Science Review, vol. 4, no. 2, p. 18, 2004.
15. A. A. Ghanbari, M. N. Kousarrizi, M. Teshnehlab, and M. Aliyari, "An evolutionary artifact rejection method for brain computer in-terface using ica," International Journal of Electrical & Computer Sciences, vol. 9, no. 9, pp. 48–53, 2009.
16. S. Siuly and Y. Li, "Improving the separability of motor imagery eeg signals using a cross correlation-based least square support vector machine for brain–computer interface," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 20, no. 4, pp. 526–538, 2012.
17. D. Fattahi, B. Nasihatkon, and R. Boostani, "A general framework to estimate spatial and spatio-spectral filters for eeg signal classifica-tion," Neurocomputing, vol. 119, pp. 165–174, 2013.
18. T. Al-Ani, D. Trad, and V. S. Somerset, "Signal processing and classification approaches for brain-computer interface," Intelligent and Biosensors, pp. 25–66, 2010.
19. L. Yu and H. Liu, "Efficient feature selection via analysis of rele-vance and redundancy," Journal of machine learning research, vol. 5, no. Oct, pp. 1205–1224, 2004.
20. B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive berlin brain–computer interface: fast acquisition of effective performance in untrained subjects," NeuroImage, vol. 37, no. 2, pp. 539–550, 2007.
21. J. Meng, L. Yao, X. Sheng, D. Zhang, and X. Zhu, "Simultaneously optimizing spatial spectral features based on mutual information for eeg classification," IEEE transactions on biomedical engineering, vol. 62, no. 1, pp. 227–240, 2014.
22. M. Senthilmurugan, M. Latha, and N. Malmurugan, "Classification in eeg-based brain computer interfaces using inverse model," Inter- national Journal of Computer Theory and Engineering, vol. 3, no. 2, p. 274, 2011.
23. N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges," Sensors, vol. 19, no. 6, p. 1423, 2019.
24. F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain–computer interfaces," Journal of neural engineering, vol. 4, no. 2, p. R1, 2007.
25. A. Mahmood, R. Zainab, R. B. Ahmad, M. Saeed, and A. M. Kamboh, "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), pp. 1034–1037, IEEE, 2017.
26. P. Bhuvaneswari and J. S. Kumar, "Support vector machine technique for eeg signals," International Journal of Computer Applications, vol. 63, no. 13, 2013.
27. H. Yang and S. Wu, "Eeg classification for bci based on csp and svm-ga," in Applied Mechanics and Materials, vol. 459, pp. 228– 231, Trans Tech Publ, 2014.
28. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," IEEE Intelligent Systems and their appli-cations, vol. 13, no. 4, pp. 18–28, 1998.
29. T. Nguyen, I. Hettiarachchi, A. Khatami, L. Gordon-Brown, C. P. Lim, and S. Nahavandi, "Classification of multi-class bci data by common spatial pattern and fuzzy system," IEEE Access, vol. 6, pp. 27873–27884, 2018.
30. Y. Li, P. P. Wen, et al., "Clustering technique-based least square sup-port vector machine for eeg signal classification," Computer methods and programs in biomedicine, vol. 104, no. 3, pp. 358–372, 2011.
31. S. Selim, M. M. Tantawi, H. A. Shedeed, and A. Badr, "A csp am-ba-svm approach for motor imagery bci system," IEEE Access, vol. 6, pp. 49192–49208, 2018.
32. C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "Bci competition 2008–graz data set a," Institute for Knowledge Dis- covery (Laboratory of Brain-Computer Interfaces), Graz University of Technology, vol. 16, 2008.
33. M. M. N. Mannan, M. A. Kamran, S. Kang, and M. Y. Jeong, "Effect of eog signal filtering on the removal of ocular artifacts and eeg-based brain-computer interface: A comprehensive study," Complexity, vol. 2018, 2018.
34. A. Liu, K. Chen, Q. Liu, Q. Ai, Y. Xie, and A. Chen, "Feature selec-tion for motor imagery eeg classification based on firefly algorithm and learning automata," Sensors, vol. 17, no. 11, p. 2576, 2017.
35. S. G. Mason and G. E. Birch, "A general framework for brain-computer interface design," IEEE transactions on neural systems and rehabilitation engineering, vol. 11, no. 1, pp. 70–85, 2003.
36. A. Schlogl, C. Brunner, R. Scherer, and A. Glatz, "20 biosig: An open-source software library for bci research," Toward brain-computer interfacing, p. 347, 2007.
37. A. Delorme and S. Makeig, "Eeglab: an open source toolbox for anal- ysis of single-trial eeg dynamics including independent component analysis," Journal of neuroscience methods, vol. 134, no. 1, pp. 9– 21, 2004.
38. M. Bhardwaj and A. K. Nadir, "Analyzing electroencephalogram signal using eeg lab," Sensors & Transducers, vol. 100, no. 1, p.51, 2009.
39. A. Bhattacharya, D. N. Bawane, and S. Nirkhi, "Brain computer interface using eeg signals," GHRCE, Nagpur.
40. K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b," Frontiers in neuroscience, vol. 6, p. 39, 2012.
41.] B. Kaur, D. Singh, and P. P. Roy, "Eeg based emotion classification mechanism in bci," Procedia computer science, vol. 132, pp. 752 758, 2018.
42. T. Joachims, "Making large-scale svm learning practical," tech. rep., Technical report, SFB 475: Komplexit¨atsreduktion in Multivariaten ,1998.
43. J. Bellary and K. R. Eddula, "Improving multi-class support vector machines training,"
44. T. Joachims, "Svmlight: Support vector machine," SVM-Light Support Vector Machine http://svmlight. joachims. org/, University of Dortmund, vol. 19, no. 4, 1999.

45. D. Bandara, J. Arata, and K. Kiguchi, "A noninvasive brain–computer interface approach for predicting motion intention of activities of daily living tasks for an upper-limb wearable robot," International Journal of Advanced Robotic Systems, vol. 15, no.2, p. 1729881418767310, 2018.

46. A. Ahangi, M. Karamnejad, N. Mohammadi, R. Ebrahimpour, and N. Bagheri, "Multiple classifier system for eeg signal classification with application to brain–computer interfaces," Neural Computing and Applications, vol. 23, no. 5, pp. 1319–1327, 2013.

47. K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," Journal of machine learning research, vol. 2, no. Dec, pp. 265–292, 2001.

48. S. Abe, Support vector machines for pattern classification, vol. 2. Springer, 2005.

49. V. Apostolidis-Afentoulis and K.-I. Lioufi, "Svm classification with linear and rbf kernels," July): 0-7. http://www. academia. edu/13811676/SVM Classification with Linear and RBF kernels.[21], 2015.

50. A. Ben-Hur, C. S. Ong, S. Sonnenburg, B. Sch¨olkopf, and G. R¨atsch, "Support vector machines and kernels for computational biology," PLoS computational biology, vol. 4, no. 10, p. e1000173, 2008.