Brain-Computer Interface Binary Classification using Regularized CSP and Stacked Concept

B. Bijitha[#1], Nandakumar Paramparambath[*2]

[#]M.Tech Student, [*]Professor
Dept. of Electronics and Communication
NSS College of engineering
Palakkad, India

***Abstract*— *Brain-Computer Interface technology is the one in which the brain signals acquired from scalp recordings are used to control external devices like artificial limbs, computers, etc. Even though studies on BCI technology are progressing, a consistent algorithm that will work with all types of data and environment are not developed so far. In this paper, an algorithm with feature extraction using regularized version of CSP and PCA, then the features are classified using the stacked concept classifier. The algorithm is evaluated using kappa coefficient and compared with existing algorithms.*

Keywords — *brain computer interfaces, EEG, feature extraction, common spatial pattern, classification, regularized linear discriminant analysis.*

I. Introduction

In day to day life, there occurs various accidents and people may lose some of their body parts, or even life at the worst. Besides this, people may become handicapped due to central nervous disorders as well. The World Report on Disability 2011 from World Health Organization and The World Bank have described that around 15% of the world population suffers from disability [1]. The disabled people find difficulty in communicating with the external world. Brain-Computer Interface (BCI) technology provides solution for disabled people (especially those suffering from disorders like cerebral palsy, amyotrophic lateral sclerosis (ALS), brainstem stroke and spinal cord injury) to help them communicate [2]. BCI has various definitions and the one defined by Wolpaw *et al.* [3] is as follows: “A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain’s normal output pathways of peripheral nerves and muscles”. BCI is sometimes known as direct brain interfaces or human computer interfaces or brain-machine interfaces (BMI). BCI has various applications like in medicine for diagnosis, environment control, entertainment, security, etc. [4, 5]. BCI can be classified into different types: exogenous and endogenous, asynchronous and synchronous [6]. BCI can also be classified as dependent and independent [3].

In BCI system the brain signals acquired are passed through signal processing stages: feature extraction/selection and the classification stage [7]. The brain signals are acquired through various brain imaging methods for example, MEG, EEG, fMRI, etc., [8]. The most commonly used method is the EEG, since it is noninvasive as well as cheaper. The brain signals acquired can be classified into mu and beta rhythms, event-related potentials, visual evoked potentials, event related synchronization/ desynchronization, and slow cortical potentials [9]. Sensorimotor rhythms are mostly exploited signal which contains oscillations in the alpha and beta frequencies (8-12 and 18-26 Hz respectively) [10]. The information contained in these signals is termed as features. In BCI system, these features needs to be extracted using various methods like CSP, PCA, ICA, wavelet transform, and autoregressive model [11]. Then using translational algorithms that may be a classifier or a regression function, these features are translated into commands [12].

Besides the significance and various applications of BCI technology, it has many problems. The most important problem to be considered is the non-stationarity of EEG signals that is the signals vary with time due to the impact of artifacts and noises due to equipments [13]. The problem mainly affects the feature extraction stage. For the classification stage, the performance is degraded due to curse of dimensionality and bias-variance trade-off. Due to these problems the efficiency of the BCI algorithm gets degraded. In this paper, the algorithm implemented is evaluated with a publicly available dataset that is described in detail below. The performance is evaluated using kappa coefficient.

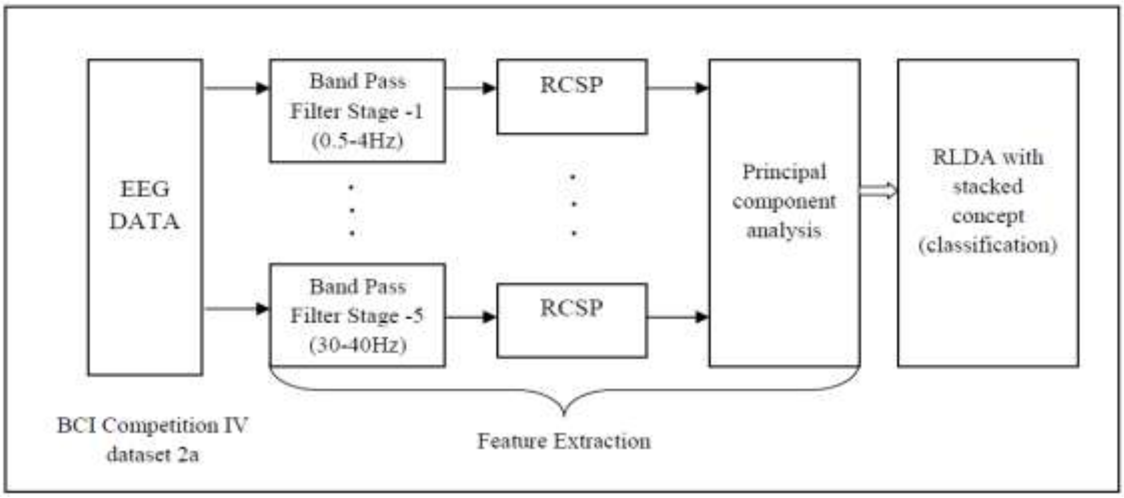

Fig.1 Block diagram of the BCI algorithm implemented in the paper

II. BCI COMPETETITION IV DATASET 2A

The dataset used here is the BCI Competition IV dataset 2a. It is provided by Graz University [14]. The dataset is a collection of EEG signals from 9 healthy subjects. The data collected from each subject contain EEG signals corresponding to four different motor imagery tasks: movement of left hand, movement of right hand, tongue movement and feet movement. The EEG signals are recorded with the help of 22 EEG channels and 3 EOG channels. The recordings were done in two different days and hence there consists of two of EEG for each subject, one for training and the other for evaluation. Each session consists of 72 trials for each movement. The signals were sampled at 250 Hz. The evaluation session consists of unknown trials and is used in the experiment to evaluate the proposed algorithm. The algorithm is evaluated using the kappa coefficient. The kappa coefficient ranges from 0(slight or no agreement) to 1 (perfect agreement). The results for the evaluation set are all declared and are available in the web page http://www.bbci.de/comp-etition/iv/results. Graz University also provides various other datasets other than dataset 2a, useful for BCI research.

III. PROPOSED METHOD

The block diagram of Fig.1 shows the signal processing stages in the proposed algorithm: multiple band-pass filtering, feature extraction using RCSP and PCA, and then classification stage using RLDA with stacked concept. The basic algorithm used in feature extraction is the CSP (common spatial pattern), and in classification linear discriminant analysis is the basic algorithm.

A. Feature Extraction Stage

*1) **Filtering***: In this stage the EEG signal from the dataset 2a is band-pass filtered such that it is split into 5 different spectral bands: 0.5-4 Hz, 4-8 Hz, 8-13 Hz, 13-30, and 30-40 Hz, according to the delta, theta, alpha, beta, and gamma rhythms respectively [15]. The concept of multiple band pass filtering is obtained from [16, 17]. FIR filters with Kaiser Window having 1 Hz transition band is set at each of the five filters. The filter is designed using 'kaiserord' MATLAB function.

2) ***Spatial Filtering:*** Spatial filtering becomes important in identifying the sources of sensorimotor rhythms corresponding to different motor imagery tasks [18]. Most commonly used spatial filtering method is the common spatial pattern (CSP) algorithm, in which the variance for one class is maximized and for other class it is minimized [19]. In CSP linear transformation of the EEG data is done using the following equation,

$$Z=W^TE, \tag{1}$$

where Z is EEG measurement after spatial filtering, W is CSP projection matrix, and E is EEG measurement [16]. In this paper, instead of CSP a regularized version, called RCSP, is used. The concept of regularization was introduced to overcome the small-sample problem in discriminant analysis [20], later it was introduced in CSP also. Fabien *et al* [21] compared different regularization methods and it was inferred from the results that RCSP outperformed CSP by 10% in classification accuracy. In the experiment the regularization parameter is introduced at two different points in CSP stage. First one is introduced in second term of CSP input parameters, wherein a fraction of first parameter is added to the second parameter. The fraction is of order 0.01 which is selected by trial and error method. Here the first and second parameters are the covariance matrices of class1 and class2 respectively for the corresponding binary class problem. The second regularization of value 0.1 (chosen by trial and error method) is introduced in the eigenvalue decomposition problem,

$$\Sigma_1 W= (\Sigma_1+0.1*\Sigma_2)WD, \tag{2}$$

where Σ_1 and Σ_2 are covariance matrices of the two classes, D is diagonal matrix containing eigen values of Σ_1 [16]. Along with regularization the spatial filtering algorithm proposed by Ang *et al* [16] is used here to enhance the CSP performance for two-class problems. Then PCA is employed. The principal component analysis is a dimension reduction method and helps in identifying the principal components [22].

B. *Classification Stage*

The basic classifier used mostly is the linear discriminant analysis (LDA) classifier [23, 24]. The basic LDA described by Vidaurre *et al* [23] is as follows:

$$V(x)= [b\ w^T][1\ x]^T \quad (3)$$

$$b= -w^T(\mu_1+\mu_2)/2,\ w=\Sigma^{-1}(\mu_1+\mu_2) \quad (4)$$

The V(x) describes the distance between a feature vector x and a hyperplane. The hyperplane is defined by the normal the bias b and the vector w. Here μ_1 and μ_2 are the sample mean of two classes and Σ, the covariance of the classes (in this case the covariance of the two classes is considered to be equal). The classification is done based on the value of V(x) that is, if the value is less than 0 then the feature vector, x, is classified as class1, otherwise as class 2 [23]. Besides as a classification algorithm, LDA has also been used as a dimensionality reduction method [25]. The classifier employed in the proposed algorithm is the regularized version of LDA. The score obtained from the first RLDA classifier is input to another RLDA classifier that has the same parameters as the first classifier. The idea of this classification technique is obtained from the concept of stacked generalization used by Nicolas *et al* [17] for classification, where 25 RLDA models per second are used in level-0 and the combined results are fed to another RLDA classifier in level-1. However in the present experiment only a single RLDA classifier is employed at each level.

IV. RESULTS

The proposed algorithm is evaluated using a publicly available BCI Competition IV Dataset 2a which is provided by the Graz University. The dataset consists of EEG recordings of 9 different persons (subject). For each subject there are two sessions taken on two different days: one for training (samples with known trials) and the other for evaluation (samples with unknown trials). The dataset has samples for four different motor imagery actions (classes): left hand, right hand, tongue movement and feet movement. There are 72 trials for each class and hence a total of 288 trials in a session. Here the proposed algorithm is implemented for binary-class problems. Since there are samples for four classes, there are six binary-lass problems: left hand v/s right hand, left hand v/s foot, left hand v/s tongue, right hand v/s foot, right hand v/s tongue, and foot v/s tongue. In the dataset there is eight second duration for each trial. From that a 1.5 second duration samples for each class are used in the experiment.

The samples are then multiple band pass filtered so that data gets discriminated on the basis of frequency. Five band pass filters are used that have transition band of 1 Hz and the cut-off frequencies are: 0.5-4 Hz, 4-8 Hz, 8-13 Hz, 13-30, and 30-40 Hz. Then the samples are passed to spatial filtering wherein the RCSP is used. There are two regularization parameters employed at this stage: one used in the second parameter for CSP that contains the covariance matrices of class2 and then fraction of covariance matrix of class1, the fraction is 0.01 and the second regularization parameter is employed in eigenvalue decomposition problem of value 0.1. The values are selected by trial and error method. The coefficients obtained are then employed to the algorithm proposed by Ang *et al* [16]. The coefficients obtained will be of dimension 25 x 25, from which the first two and last columns are selected. Hence a total 20 features will be obtained (5*4). Then the principal component analysis is performed. The PCA helps in reducing the noise factors. The number of trials for binary classification is 144, but in the experiment only 72 trials are taken (36 trials from each of the two classes considered). The 72 trials with 20 features are then applied to classification stage, wherein the features will be fed to two classifiers with same parameters. The features fed to first classifier, yields score which is then fed to next classifier. The classifier employed here is the RLDA classifier. The regularization parameter value is 0.1, which is again selected by trial and error method. Then kappa values are estimated. The algorithm is first examined by 10-fold cross validation, for which the results are satisfactory. Then the algorithm is tested with the training data itself. Then the algorithm is evaluated with the evaluation data. The kappa values obtained are listed in the tables. The cross validation values are written inside simple brackets and that with the training data inside square brackets. The obtained kappa values are compared with the algorithm proposed by Nicolas *et al* [17]. The values that are high for algorithm used in the experiment are shown in bold. The results show that the proposed algorithm outperforms the SRLDA [17] in some cases. In the binary problem the propose algorithm outperforms the SRLDA [17] for five subjects. In rest all cases the algorithm outperforms the SRLDA [17] for two-four subjects. However the average results become less since for certain subjects the kappa values obtained are low. Even though the results shows only a slight improvement in certain cases, while considering the number of samples taken (72 trials with 20 features), the algorithm proves to be efficient. The experiment is implemented in MATLAB software and with the help of BIOSIG toolbox.

TABLE I.
KAPPA VALUES OBTAINED FOR SIX BINARY-CLASS PROBLEMS WITH PROPOSED ALGORITHM AND THE TABLE ALSO SHOWS COMPARISON WITH SRLDA [17].

Binary Class Problem	Method	Subjects									Avg.
		A1	A2	A3	A4	A5	A6	A7	A8	A9	
Left-Right	SRLDA [17]	0.82	0.39	0.92	0.51	0.89	0.49	0.96	0.96	0.81	0.75
	Algorithm with RCSP and Stacked Concept	0.17	**0.53**	**0.97**	0.48	0.63	0.26	0.90	0.75	0.62	0.59
		(0.96)	(0.88)	(0.97)	(0.96)	(0.94)	(0.91)	(0.87)	(0.97)	(0.94)	(0.93)
		[0.88]	[0.88]	[0.86]	[0.83]	[0.91]	[0.75]	[0.94]	[0.94]	[0.88]	[0.87]
Left-Foot	SRLDA [17]	0.96	0.82	0.96	0.75	0.71	0.61	1	0.88	0.96	0.85
	Algorithm with RCSP and Stacked Concept	0.06	0	0.68	**0.89**	0.19	**0.96**	0.19	**1**	0.07	0.45
		(0.89)	(1)	(0.88)	(0.85)	(0.95)	(0.87)	(0.87)	(0.84)	(1)	(0.90)
		[0.86]	[0.80]	[0.83]	[0.86]	[0.75]	[0.86]	[0.83]	[0.86]	[0.83]	[0.83]
Left-Tongue	SRLDA [17]	0.93	0.63	0.90	0.81	0.68	0.33	0.96	0.92	0.94	0.79
	Algorithm with RCSP and Stacked Concept	**0.96**	0.01	0.50	0.79	0.26	**0.72**	0.12	**1**	**0.97**	0.59
		(0.91)	(0.90)	(0.87)	(0.92)	(0.92)	(0.95)	(0.94)	(0.93)	(0.97)	(0.92)
		[0.94]	[0.61]	[0.61]	[0.94]	[0.72]	[0.75]	[0.88]	[0.91]	[0.88]	[0.80]
Right-Foot	SRLDA [17]	0.97	0.92	0.94	0.90	0.76	0.56	0.99	0.88	0.64	0.84
	Algorithm with RCSP and Stacked Concept	0.91	0.18	0.33	0.44	0.74	**0.64**	0.19	0.84	**0.84**	0.56
		(0.85)	(0.76)	(0.91)	(0.94)	(1)	(0.80)	(0.96)	(0.97)	(0.96)	(0.96)
		[0.72]	[0.86]	[0.86]	[0.83]	[0.77]	[0.83]	[0.77]	[0.88]	[0.83]	[0.81]
Right-tongue	SRLDA [17]	0.99	0.53	0.94	0.86	0.86	0.36	1	0.75	0.83	0.79
	Algorithm with RCSP and Stacked Concept	0.61	**1**	0.62	0.79	0.07	**0.46**	0.97	0.56	**0.83**	0.65
		(0.91)	(0.83)	(0.97)	(0.97)	(0.93)	(0.88)	(0.97)	(0.96)	(1)	(0.93)
		[0.94]	[0.97]	[0.91]	[0.86]	[0.69]	[0.80]	[0.97]	[0.88]	[0.88]	[0.87]
Foot-Tongue	SRLDA [17]	0.72	0.81	0.86	0.61	0.43	0.68	0.75	0.81	0.83	0.72
	Algorithm with RCSP and Stacked Concept	**1**	0.20	0.55	0.12	**0.48**	**0.74**	**0.97**	**0.97**	0.41	0.60
		(1)	(0.83)	(0.88)	(0.84)	(0.86)	(1)	(0.93)	(0.97)	(0.97)	(0.92)
		[0.83]	[0.80]	[0.77]	[0.63]	[0.86]	[0.86]	[0.94]	[0.91]	[0.88]	[0.83]

V. CONCLUSIONS AND FUTURE WORK

The proposed algorithm has three main stages, multiple band pass filtering, spatial filtering and then the classification. The regularization is introduced at two signal processing stages: at feature extraction (spatial filtering) and then at the classification stage. The efficiency of the algorithm is evaluated using kappa coefficient. At first the algorithm is evaluated using 10-fold cross-validation and the obtained results are high. Then the algorithm is evaluated using evaluation data set, and the results compared with that of Nicolas *et al* [17]. The results outperform the one proposed by Nicolas *et al* [17] in certain cases only. However while considering the fact that the features used in the proposed algorithm is only 20 and the number of trials used is 72, whereas in [17], the authors have used 144 trials, the results obtained are remarkable. The algorithm has large scope for improvement. The algorithm is implemented for binary class problems only, hence can be extended for four-class classification. Also the algorithm is tested with only one dataset, it can also be tested with various other datasets.

REFERENCES

[1] *World Report on Disability* 2011 from World Health Organization and The World Bank available at http://www .who.int/disabilities/world_report/2011/report.pdf

[2] Jonathan R. Wolpaw, "Brain-Computer Interface Research Comes of Age: Traditional Assumptions Meet Emerging Realities", *Journal of Motor Behavior*, Vol. 42, No. 6, pp. 351-353, 2010.

[3] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, Theresa M. Vaughan, "Brain-computer interfaces for communication and control", *Elsevier Science Ireland Ltd*, Clinical Neurophysiology 113, pp. 767–791, 2002.

[4] Cincotti. F, Mattia. D, Aloise. F, Bufalari. S, Schalk. G, Oriolo. G, Cherubini. A, Marciani. M.G, Babiloni. F, "Non-invasive brain-computer interface system: Towards its application as assistive technology", *Robotics and Neuroscience*, Vol. 75, pp. 796-803, April 2008.

[5] Roman. K, Benjamin. B, Gabriel. C, Klaus-Robert. M, "The Berlin Brain-Computer Interface (BBCI)-Towards a new communication channel for online control in gaming applications", *Multimedia Tools and Applications* 33, pp. 73–90, 2007.

[6] Aboul Ella Hussanien, and Ahmed Taher Azar," Brain Computer Interface Current Trends and Applications", Intelligent Systems Reference Library, *Vol. 74, Springer Interational Publishing*, 2015.

[7] Obed Carrera-Leon, Juan Manuel Ramirez, Vicente Alarcon-Aquino, Mary Baker, David DCroz-Baron, Pilar Gomez-Gil, "A Motor Imagery BCI Experiment using Wavelet Analysis and Spatial Patterns Feature Extraction", *IEEE Conference on Engineering Applications (WEA)*, 2012 Workshop on, 2012.

[8] Katarzyna Blinowska, Piotr Durka, "Electroencephalography (EEG)", *Wiley Encyclopedia of Biomedical Engineering*, John Wiley & Sons, Inc., 2006.

[9] Yijun Wang, Xiaorong Gao, Bo Hong, Chuan Jia, and Shangkai Gao, "Brain-Computer Interfaces Based on Visual Evoked Potentials Feasibility of Practical System Designs", *IEEE Engineering in Medicine and Biology Magazine*, pp.64-71, September/October, 2008.

[10] J. R. Wolpaw, D. J. Mcfarland, and T. M. Vaughan, "Brain–Computer Interface Research at The Wadsworth Center", *IEEE Transactions on Rehabilitation Engineering*, Vol. 8, No.2, pp.222-226, June 2000.

[11] Sun-Yuge, Ye-Ning, Zhao-Lihong, Xu-Xinh, "Research on Feature Extraction Algorithms in BCI*", IEEE, Chinese Control and Decision Conference (CCDC)*, pp.5874-587, 2009.

[12] Dennis J. McFarland, Dean J. Krusienski and Jonathan R. Wolpaw, "Brain computer interface signal processing at the Wadsworth Center: mu and sensorimotor beta rhythms", *Progress in Brain Research*, Vol. 159, pp.411-419, February 2006.

[13] P.Shenoy, M.Krauledat, B.Blankertz, R.P.N.Rao, and K.R.Müller, "Towards Adaptive Classification for BCI," *Journal of Neural Engineering*, Vol. 3, pp. R13, March 2006.

[14] M. Tangermann et al., "Review of the BCI competition IV," Frontal Neuroscience, Vol. 6, pp. 55, July 2012.

[15] L.F.Nicolas-Alonso and J.Gomez Gil,"Brain computer interfaces, a review," *Sensors*, vol.12, pp.1211–1279, Jan.2012.

[16] 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," *Frontal Neuroscience*, Vol. 6, March 2012.

[17] Luis F.Nicolas-Alonso, Rebeca Corralejo, Javier Gomez-Pilar, Daniel Álvarez ,and Roberto Hornero, "Adaptive Stacked Generalization for Multiclass Motor Imagery-Based Brain Computer Interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*,Vol.23, No.4, pp. 702-712, July2015

[18] Wojciech Samek, Motoaki Kawanabe, and Klaus-Robert Muller , "Divergence-Based Framework for Common Spatial Patterns Algorithms", *IEEE Reviews in Biomedical Engineering*, Vol. 7, pp.50-72, 2014.

[19] B. Blankertz, R.Tomioka, S.L.M.Kawanabe, and K.R.Muller, "Optimizing Spatial Filters for Robust EEG Single-Trial Analysis", *IEEE Signal Processing Magazine*, Vol. 25, pp. 41–56, January 2008.

[20] Haiping Lu, How-Lung Eng, Cuntai Guan, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos, "Regularized Common Spatial Pattern with Aggregation for EEG Classification in Small-Sample Setting", *IEEE Transactions on Biomedical Engineering*, Vol. 57, No. 12, pp.2936-2946, December 2010.

[21] Fabien Lotte, and Cuntai Guan, "Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms*", IEEE Transactions on Biomedical Engineering*, Vol. 58, No. 2, pp.355-362, February 2011.

[22] Li Ke, Rui Li, "Classification of EEG Signals by Multi-Scale Filtering and PCA", *IEEE*, pp. 362-366, 2009.

[23] C.Vidaurre, M.Kawanabe, P.vonBunau, B.Blankertz, and K.Muller, "Toward Unsupervised Adaptation of LDA for brain–computer interfaces", *IEEE Transaction on Biomedical Engineering*, Vol.58, pp.587–597, March 2011.

[24] Z. Jinyin, G. Sudre, L. Xin, W. Wei, D. J. Weber, and A. Bagic, "Clustering Linear Discriminant Analysis for MEG-Based Brain Computer Interfaces", *IEEE Transaction on Neural System and Rehabilitation Engineering*, Vol.19, pp. 221–231, March 2011.

[25] D. Cai, X. He, and J. Han, "SRDA: An efficient Algorithm for Large Scale Discriminant Analysis," *IEEE Transaction on Knowledge Data Engineering*, Vol. 20, pp. 1–12, January 2008.