

Feature Extraction of EEG based Motor Imagery Using CSP based on Logarithmic Band Power, Entropy and Energy

Majid Aljalal, Ridha Djemal, Khalil AlSharabi and Sutrisno Ibrahim

Electrical Engineering Department
College of Engineering, King Saud University
P.O.Box 800 – 11421 Riyadh, KSA

algalalmajid@hotmail.com, {rdjemal, kabdulghani, suibrahim}@ksu.edu.sa

Abstract— One of the most important applications of brain-computer interface (BCI) is assisting disabled people to control an external device by using motor imagery (MI). We focused in this paper on the classification of two types of MI tasks (left-hand \times right-hand and right-hand \times foot) in the electroencephalogram (EEG) signal. We compare various feature extraction techniques by combining common spatial pattern (CSP) with several features: variance, energy, entropy and logarithmic band power (LBP). Three types of classifiers were employed for classification: linear discriminant analysis (LDA), support vector machines (SVM) and artificial neural network (ANN). We tested our proposed method using data recorded from 17 subjects, provided by BCI-Competition III and IV. The results show that features extracted using a combination of CSP and LBP produce highest classification accuracy. LDA is more suitable than other classifiers to classify features extracted using CSP and LBP.

Keywords— artificial neural network, band power, brain-computer interface, common spatial pattern, energy, entropy, motor imagery, linear discriminant analysis, support vector machines.

I. INTRODUCTION

Brain-computer interface (BCI) is a system interface to exchange messages and commands between user's brain and a computer or an external device. Providing a possibility to drive an external device directly from the brain is the major objective of BCI [1]. This will be very valuable to assist people with body disability. BCI can be mostly classified into invasive and noninvasive models. In invasive BCI, a surgery will be needed to place the electrodes inside the brain to capture the signals. While in non-invasive the electrodes are placed outside the brain. The signals are stronger in the invasive method but it is less practical because of the need for surgery [2]. Brain activities are sometimes called "rhythms". These rhythms are microvolt electric signals which are created in our brain while we are doing a task [3, 4].

Electroencephalogram (EEG) signals can be used to capture brain activity. There are three main types of EEG signals used in the BCI system: P300, SSVEP and ERD/ERS. P300 is a localized response to a joined visual, auditory, or tactile stimulus, and it is mainly read from the parietal lobe

during 300ms after starting of the stimulus [5, 6]. Steady-state visual evoked potentials (SSVEP) is a response to a visual stimulus modulated at a frequency greater than 6 Hz, which can be read from the occipital area [7, 8]. The event-related desynchronization (ERD) and event-related synchronization (ERS) are prompted by performing mental tasks, such as motor imagery, mental arithmetic, or mental rotation [9]. In the case of the motor imagery paradigm, the mu (8–13Hz) and beta (14–30Hz) rhythms of the sensorimotor cortex are used [10]. During physical and motor imagery of right and left hand movements, beta band brain activation β -ERD happens predominantly over the contralateral left and right motor areas and a β -ERS ipsilaterally. The post movement ERS related to ceasing to move can also be found over the contralateral motor areas. The appropriate processing (filtering, features extraction and classification) of MI states is important to generate appropriate commands to control rehabilitation devices or other applications.

In motor imagery, selection a suitable technique for features extraction is the most important point to extract important information in the signal to recognize what human think. One of the most widely used algorithm for features extraction is common spatial patterns (CSP). In 2000, CSP started to be employed to detect the event-related desynchronizations [11]. A. Baziyad et al. [12] aggregated the features extracted from CSP with others extracted from wavelet transform (WT). They concluded the best results were gotten when using SVM classifier with an average classification accuracy of 75% over three subjects. R. Chatterjee et al. [13] employed wavelet-based energy-entropy & RMS with SVM and MLP classifiers and obtained accuracy of 85% and 85.71% respectively over only one subject. Most researchers that use CSP employed variance value to construct the features vector. This is because that CSP gives peak variances for the discrimination of two classes of EEG related motor imagery [14]. In this paper, we deployed various feature extraction techniques on the EEG data using common spatial pattern based on entropy, energy and logarithmic band power. Three types of algorithms were employed for classification: LDA, SVM and ANN. The aim of this paper is to compare them and to recommend a suitable technique to be used with

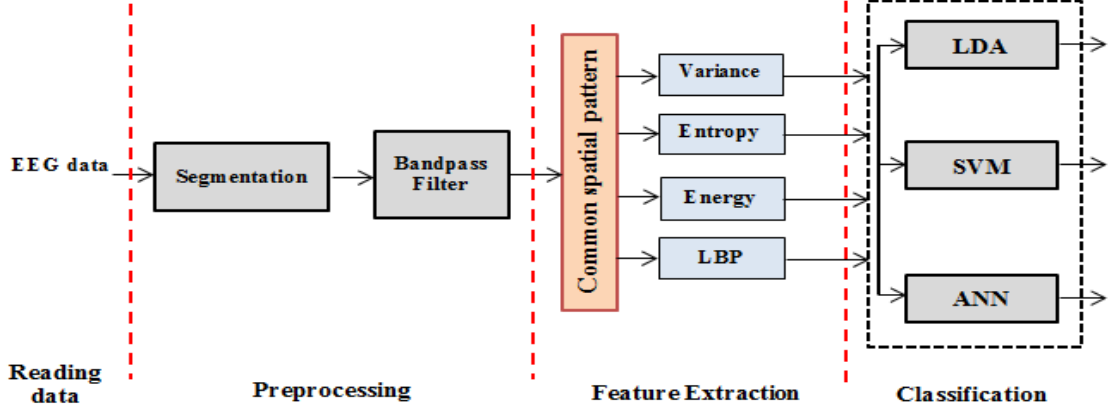


Fig.1 Block diagram of the proposed methods.

CSP for extracting features of synchronous two-class-imagery-based brain-computer interface experiments. Classification accuracy with cross-validation over 17 subjects was employed to evaluate the performance of our proposed approaches.

The remainder of this paper is organized into three sections. Section II describes the EEG datasets used in our work and briefly explains various features extraction and classification techniques. Experimental results are presented and discussed in Section III. Finally, the paper is concluded in Section IV.

II. METHODS

In this section we describe the methods of feature extraction and classification techniques as well as their validation using Matlab simulation tools. Fig. 1 shows the block diagram of the proposed methods. Firstly, EEG data were read and then segmented into time windows of 3s (and 2.5s) with offset of 0.5s. Then, the output of the segmentation process was fed into bandpass filter. In this experiment, we employed several frequency bands to validate the proposed approaches. Next, common spatial pattern algorithm was applied then to the filtered signals. In order to extract the feature vectors and to improve the classification accuracy, we used variance, entropy, energy and logarithmic band-power with common spatial pattern. Linear discriminant analysis, support vector machines and artificial neural network techniques were then employed as classifiers. All possible combinations of the proposed approaches were implemented and verified.

A. Datasets

In our work, we used three groups of data sets from BCI competition (III and IV) including 17 subjects. The data (III-IVa) [15] were provided by Berlin BCI group: Fraunhofer First, Intelligent Data Analysis Group (Klaus Robert Muller, Benjamin Blankertz) and campus Benjamin Franklin of the Charite University Medicine Berlin. These data are for two class motor imagery (right-hand and foot) taken from 5

subjects. For each subject they conducted 280 (right-hand 140 / foot 140) trials. The data were recorded from 118 EEG channels with 100Hz sampling rate for each trial. The second data set is III-IIIa. These datasets are provided by Graz University of Technology. These data set is recorded when three subjects were trying to perform four different types of motor imagery tasks (left-hand “LH”, right-hand “RH”, Feet “F” and Tongue “T”). In our work, only two tasks (right-hand and left-hand) were used. The recording was achieved by a 64-channel EEG amplifier from Neuroscan. The EEG signals were sampled with 250 Hz, band-pass filtered between 1 and 50 Hz, and notch filtered at 50 Hz to suppress the noise [15]. The third data set is IV-IIa [16]. This is a four class dataset. This was also provided by Graz University of Technology. They were recorded from nine subjects. This includes; Class-I (right-hand), Class-II (left-hand), Class-III (both feet), and Class-IV (tongue). In our work, only two tasks (right-hand 144 / left-hand 144) were used. The EEG signals were captured by 25 electrodes with 250Hz sampling rate.

B. Preprocessing

For analysis raw EEG, data were arranged into E-matrices. Each E-matrix is of size $N \times T$, where N is the number of channels and T is the number of EEG samples per channel in a specific interval of time. In our work, for each trial, time windows of 3s with offsets of 0.5s were applied for all subjects except subjects of IV-IIa, we applied time windows of 2.5s.

After windowing raw signal, EEG data were filtered using bandpass Butterworth filter of order 4 to remove the artifacts coming from other frequencies except for mu and beta rhythms [17]. In this experiment, several frequency bands were applied for all subjects and all approaches.

C. Feature Extraction

There are several types of feature extraction techniques used for two class imagery task discrimination. In this experiment, we have used the most popular and widespread technology. This method is common spatial pattern (CSP)

based on variance. In this work, we also proposed to combine CSP with entropy, energy and logarithmic band power.

1. Common spatial pattern

In this study, all channels were used without Laplacian filter. We used common spatial pattern algorithm as spatial filter that lead to peak variances for the discrimination of two classes of EEG related right-hand and left-hand (or foot) motor imagery and to reduce the number of channels used [14]. Computing the projection matrix constructs a set of common spatial pattern filters. The algorithm starts by computing normalized spatial co-variance for both classes. This is achieved by the following equations,

$$C_{CI} = \frac{E_{CI} E_{CI}'}{\text{trace}(E_{CI} E_{CI}')} , C_{CII} = \frac{E_{CII} E_{CII}'}{\text{trace}(E_{CII} E_{CII}')} \quad (1)$$

where E_{CI} and E_{CII} denote two single trials, under two conditions (class I and class II), of size $ch \times T$, where ch is the number of channels and T is the number of samples per channel. E' is the transpose of the E and $\text{trace}(EE')$ is the sum of diagonal elements of EE' . Then, the averaged normalized covariance \overline{C}_{CI} and \overline{C}_{CII} are calculated by averaging over all the trials of each class. The overall composite spatial covariance is given by

$$C_C = \overline{C}_{CI} + \overline{C}_{CII} \quad (2)$$

And factorized into eigenvalues and eigenvectors such as

$$C_C = U_C \lambda_C U_C' \quad (3)$$

where U_C is the matrix of eigenvectors and λ_C is the diagonal matrix of eigenvalues. All the eigenvalues are arranged in the descending order. After this, the whitening transformation P is computed such as

$$P = \sqrt{\lambda_C^{-1}} U_C' \quad (4)$$

After this we have to find

$$S_{CI} = P \overline{C}_{CI} P' \text{ and } S_{CII} = P \overline{C}_{CII} P' \quad (5)$$

In order to test these calculations the sum of corresponding eigenvalues of S_{CI} and S_{CII} should be identity matrix and S_{CI} and S_{CII} have the same eigenvectors such as

$$S_{CI} = B \lambda_{CI} B' \quad (6)$$

$$S_{CII} = B \lambda_{CII} B' \quad (7)$$

$$\lambda_{CI} + \lambda_{CII} = I \quad (8)$$

where B is any orthonormal matrix which satisfy

$$B'(S_{CI} + S_{CII})B = I \quad (9)$$

The largest eigenvalues with the corresponding eigenvectors for S_{CI} have the smallest eigenvalues for S_{CII} and vice versa. This indicates that eigenvalues for one class will be

maximized at a point whereas the other class will have eigenvalues minimum at that same point. Thus, the covariance between the two classes is successfully maximized. A set of CSP filters (projection matrix) can be obtained as

$$W_{CSP} = P' B = [w_1 \ w_2 \ \dots \ w_{ch-1} \ w_{ch}] \in R^{ch \times ch} \quad (10)$$

First CSP filter w_1 provides the maximum variance of class I, and the last CSP filter w_{ch} provides the maximum variance of class II. In our work, we selected first and last m filters such as

$$W_{CSP} = [w_1 \ w_2 \ \dots \ w_m \ w_{ch-m+1} \ w_{ch-m} \ \dots \ w_{ch}] \in R^{2m \times ch} \quad (11)$$

The filtered signal $s(t)$ is given by

$$s(t) = W_{CSP} e(t) = [s_1(t) \ s_2(t) \ \dots \ s_d(t)]' \quad (12)$$

where $e(t)$ is the signal to be reduced, d is called reduction number and equal to $2m$. The reduction number is the number of channels that desired to reduce. Thus, for each class EEG sample matrix we are going to select only small number of signals m that are most important for discrimination between the two classes. The feature vectors $f = (f_1, f_2, f_3, \dots, f_{2m})'$ can be calculated by the following equation;

$$f_i = \log \left[\frac{\text{var}[s_i(t)]}{\sum_{i=1}^{2m} \text{var}[s_i(t)]} \right] \quad (13)$$

Thus, d features were obtained for each trial as a result of common spatial filtering.

As our proposition, the features vectors were formed also using the following techniques:

Let $S(n)$, and $n = 1, 2, \dots, N$, is a discrete signal, and N is sample number of the signal, then

Entropy Features:

None normalized Shannon entropy -

$$Ent = \sum_{n=1}^N |S(n)|^2 \log |S(n)|^2 \quad (14)$$

One feature is obtained for each channel. Thus, d features were obtained for each trial as a result of CSP-entropy.

Energy Features:

The energy of the signal-

$$Eng = \sum_{n=1}^N |S(n)|^2 \quad (15)$$

Similarly, the number of features was d as obtained as a result of CSP-energy.

Logarithmic band power (LBP) Features:

Logarithmic band power of the signal-

$$Eng = \log\left(\frac{1}{N} \sum_{n=1}^N |S(n)|^2\right) \quad (16)$$

As a result of CSP-LBP, d features were obtained.

Table I includes features vectors size for all proposed approaches. The size depends on the number of trials of each group of data sets and selected reduction number (d). As d increase, the feature vector size will also increase.

TABLE I. SIZE AND NUMBER OF FEATURE VECTORS FOR ALL APPROACHES.

Feature vectors	Size		
	(No. of trials x No. of Features per channel x reduction number)		
	III-IIa	III-IVa	IV-IIa
CSP+Var	$90 \times 1 \times d$	$280 \times 1 \times d$	$144 \times 1 \times d$
CSP+Ent	$90 \times 1 \times d$	$280 \times 1 \times d$	$144 \times 1 \times d$
CSP+Eng	$90 \times 1 \times d$	$280 \times 1 \times d$	$144 \times 1 \times d$
CSP+LBP	$90 \times 1 \times d$	$280 \times 1 \times d$	$144 \times 1 \times d$

CSP: common spatial pattern; Var: variance; Ent: entropy; Eng: energy; LBP: logarithmic band power

D. Classification and cross-validation

There are several algorithms used for classification. In this experiment, we have employed the most popularly used types of classification techniques to classify the obtained feature, Linear Discriminant Analysis (LDA) [18], Support Vector Machine (SVM) [19], and Artificial Neural Network (ANN). The aim was to compare them and to conclude which one provides best results for two motor imagery tasks classification. LDA is based on mean vectors and covariance matrices of feature vectors for individual classes. It uses a hyperplane to differentiate between classes, reducing the variance within the class and make the most of the variance between the classes [18].

In order to make linear SVM convergence fast, parameter 'BoxConstraint' was reduced from 0.5 to 0.1. For ANN, we design the network with one input layer, one hidden layer and one output layer using Matlab. The hidden layer was designed with 5 nodes and output layer with 2 nodes. However, the nodes number of input layer was depended on the number of obtained features in the vectors. We have changed the transfer functions of hidden layer and output layer to 'satlins' and 'tansig', respectively, the train function to 'trainbr', train learning to 0.01 and others parameters as well. All these modifications were done to obtain better results.

In our experiment, we used k -fold cross-validation technique to obtain the classification accuracy. In this technique, the dataset is arbitrarily separated into k equal parts (k subsets) [20]. All the subsets are employed for training except one for the test (validation). The cross-validation is recurrent k times (folds). Then the results of k times are averaged to produce a single classification rate. In this

experiment, we used 5-fold cross validation (20% for testing and 80% for training) for datasets III-IIIa and III-IVa. For data set IV-IIa, we used 6-fold cross validation (17% for testing and 83% for training) to avoid the remainder when splitting the trials subsets ($144/5 = 28.8$ subsets).

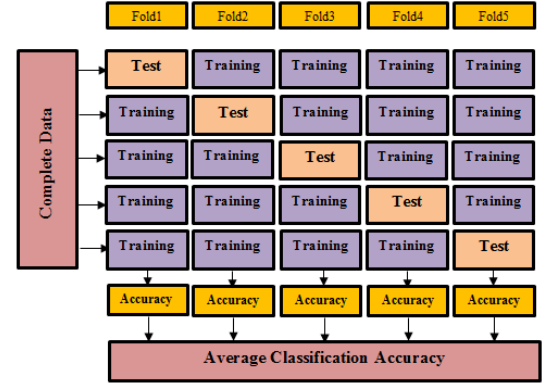


Fig.2 Cross-validation methodology for 5-cross-validation

Figure 2 shows the procedure of the cross-validation procedure; including how the training and test subsets were classified for one subject. The classification accuracy can be given by

$$accuracy = \left(\frac{N_{correct}}{N_{total}} \right) \times 100\% \quad (17)$$

where N_{total} is the overall number of vectors to be classified and $N_{correct}$ is the number of the correct vectors.

III. RESULTS AND DISCUSSION

As already mentioned, dataset (III-IVa) has five subjects with 280 trials for each one divided for training (224 trials) and testing (56 trials) data according to 5-fold cross-validation. This dataset includes two mental tasks, right-hand and right-foot. As well dataset (III-IIIa) has three subjects with 90, 60, and 48 trials, respectively. Similarly, dataset (IV-IIa) has nine subjects with 144 trials for each one divided for training (120) and testing (24). Before feature extraction, the signals were segmented to time window of 3s and 2.5s for datasets III (IIIa-IVa) and III-IVa, respectively, with offset of 0.5 s. Figure 3 shows a 2D plot of randomly selected feature vectors of IIIa dataset at 8-34Hz frequency band. It is clear from the figure that the features extracted by CSP and LBP can be easily classified. Table II shows classification accuracy of each subject for all proposed features and classified using LDA. In order to show the effect of bandpass filtering, the table also shows the average classification accuracies with several frequency bands. It can be noticed that the best results were obtained when using logarithmic band power over all frequencies bands. For all approaches, reduction number was set to 2 ($d=2$). In an instance, at 8-34Hz frequency band, CSP-based LBP provided better results than other approaches with average classification accuracy, over 17 subjects, of 80.08%.

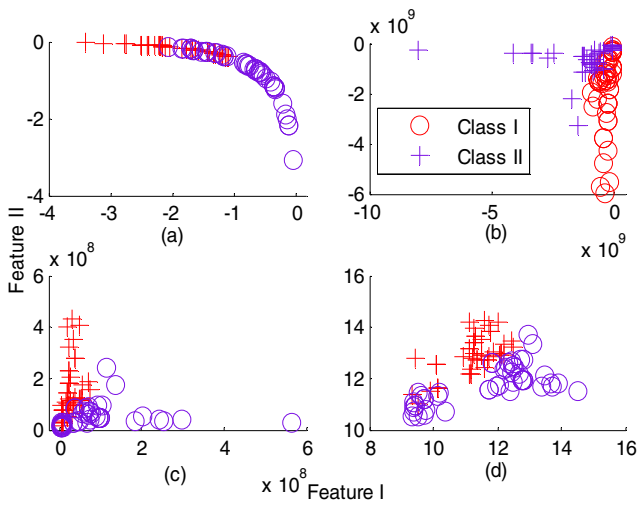


Fig. 3: 2D plot of randomly selected feature vectors of subjects of IIIa at 8-34Hz frequency band using (a) CSP + variance, (b) CSP+ Entropy, (c) CSP+ Energy, and (d) CSP + logarithmic band power.

Table III shows the average classification accuracies of CSP features filtered at 8-34Hz with different values of dimensions. It can be noticed that the results were a little improved especially at $d=12$ with average classification accuracy of 80.59%. In spite of increasing reduction number

(d), CSP based LBP still provides the highest average classification accuracy over 17 subjects. Figure 4 shows the average accuracy for all approaches classified by LDA, SVM and ANN classifiers. It can be noticed that LDA provide the best results with logarithmic band power features. ANN provided the worst results with CSP-LBP. The highest average classification accuracy obtained from ANN classifier was with CSP-based variance with average accuracy of 78.08%.

IV. CONCLUSION AND FUTURE STUDY

In our work, we have focused on the classification of two types of motor imagery tasks. We used common spatial pattern (CSP) which originally depends on variance to form feature vectors. In our work, we proposed to use other techniques in addition to variance: energy, entropy and band power. Three types of classifiers were employed for classification process: linear discriminant analysis (LDA), support vector machines (SVM) and artificial neural network (ANN). In the preprocessing, we investigated several bandpass Butterworth filters to compare the performance of the proposed methods in the different frequency bands. We observe that CSP based on logarithmic band power is more appropriate for features extraction than CSP based on entropy, energy or variance.

TABLE II
AVERAGE CLASSIFICATION ACCURACY FOR ALL PROPOSED APPROACHES AT DIFFERENT FREQUENCY BANDS

		BCI competition III								BCI competition IV									
		Data set IIIa			Data set IVa					Data set IIa									
Subject		S31	S32	S33	S51	S52	S53	S54	S55	S91	S92	S93	S94	S95	S96	S97	S98	S99	Avg.
8-34	CSP+Var	97.78	56.67	81.67	79.64	94.29	54.64	86.79	88.93	88.89	56.94	96.53	72.92	61.11	61.11	63.89	97.22	89.58	78.15
	CSP+Eng	95.56	66.67	83.33	83.21	91.79	52.50	92.50	88.93	81.94	61.11	95.83	72.22	56.94	59.72	70.14	95.83	82.64	78.29
	CSP+Ent	95.56	68.33	83.33	83.93	91.79	51.79	92.50	89.64	83.33	60.42	95.14	73.61	56.94	59.72	70.14	95.83	82.64	78.51
	CSP+LBP	98.89	66.67	88.33	81.07	97.50	57.14	92.14	89.29	87.50	55.56	95.83	75.00	61.81	59.03	70.83	96.53	88.19	80.08
8-30	CSP+Var	96.67	53.33	81.67	79.29	94.29	54.64	86.79	89.29	88.19	56.25	96.53	71.53	50.00	61.81	73.61	97.22	88.89	77.65
	CSP+Eng	94.44	65.00	83.33	83.21	91.79	49.29	92.86	88.93	80.56	59.03	95.83	73.61	42.36	58.33	72.92	95.83	82.64	77.06
	CSP+Ent	94.44	65.00	83.33	83.57	91.79	49.64	92.86	90.00	81.94	59.72	95.83	74.31	41.67	58.33	73.61	95.83	82.64	77.32
	CSP+LBP	98.89	65.00	90.00	82.14	97.50	55.36	92.14	88.93	86.81	54.86	96.53	73.61	43.75	58.33	73.61	97.22	87.50	78.95
8-32	CSP+Var	96.67	58.33	81.67	80.36	94.29	54.29	86.79	89.29	88.19	54.86	96.53	71.53	57.64	61.11	70.14	97.22	89.58	78.15
	CSP+Eng	94.44	68.33	83.33	83.21	91.79	51.43	92.50	88.93	80.56	60.42	95.83	73.61	52.08	59.03	72.22	95.83	82.64	78.01
	CSP+Ent	94.44	70.00	83.33	83.57	91.79	50.71	92.86	89.64	83.33	60.42	95.83	74.31	52.78	59.03	72.22	95.83	82.64	78.40
	CSP+LBP	97.78	66.67	88.33	81.79	97.50	56.07	92.14	88.93	87.50	55.56	95.83	74.31	56.94	58.33	72.92	96.53	88.19	79.72
10-34	CSP+Var	98.89	61.67	81.67	82.14	94.29	56.07	90.71	88.93	90.97	58.33	95.14	74.31	63.89	61.81	64.58	97.22	86.81	79.26
	CSP+Eng	96.67	61.67	81.67	84.29	91.43	51.43	94.64	88.93	86.11	56.94	95.14	68.06	62.50	57.64	72.22	97.22	85.42	78.35
	CSP+Ent	96.67	61.67	83.33	85.00	91.43	51.43	94.64	88.93	86.11	57.64	95.14	68.75	62.50	59.03	72.92	97.22	85.42	78.70
	CSP+LBP	98.89	66.67	88.33	83.21	97.14	54.29	95.36	90.36	92.36	52.78	96.53	71.53	64.58	60.42	72.22	97.92	87.50	80.59
8-25	CSP+Var	96.67	55.00	80.00	78.57	93.93	53.93	87.14	88.93	85.42	59.03	97.22	72.22	42.36	56.94	72.22	96.53	88.89	76.76
	CSP+Eng	91.11	63.33	83.33	82.50	91.79	49.64	92.86	88.57	79.86	60.42	95.83	72.22	34.72	57.64	75.00	95.83	82.64	76.31
	CSP+Ent	92.22	65.00	83.33	82.86	91.79	49.29	93.57	90.00	80.56	60.42	95.83	72.22	35.42	56.94	72.92	95.83	82.64	76.52
	CSP+LBP	97.78	60.00	90.00	81.79	97.50	55.36	92.50	89.29	86.11	59.03	96.53	72.22	35.42	56.94	75.69	97.22	87.50	78.29

TABLE III
EFFECT OF REDUCTION NUMBER ON THE CLASSIFICATION ACCURACY AT FREQUENCY BAND OF 8-34Hz

	Reduction Number										
	$d=2$	$d=4$	$d=6$	$d=8$	$d=10$	$d=12$	$d=14$	$d=16$	$d=18$	$d=20$	$d=22$
CSP+Var	78.15	78.49	78.27	79.72	80.18	79.76	79.34	78.92	78.83	78.21	78.03
CSP+Eng	78.29	77.55	77.36	77.95	78.27	78.79	78.14	78.16	77.33	76.69	76.15
CSP+Ent	78.51	77.51	77.48	78.62	78.65	78.87	78.44	78.46	77.70	76.75	76.46
CSP+LBP	80.08	78.87	78.89	80.18	80.56	80.59	80.29	79.54	79.05	78.51	78.99

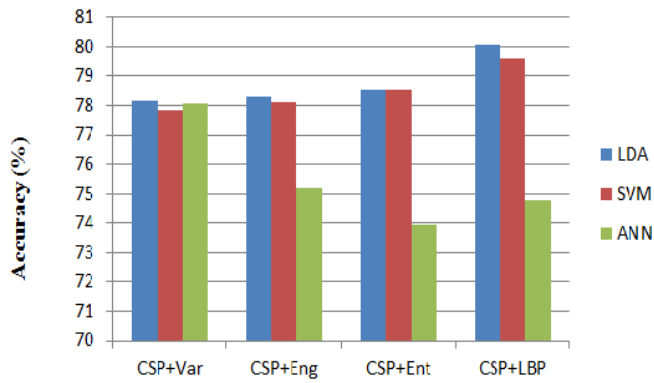


Fig. 4: Average accuracy of classification (%) of all proposed approaches using LDA, SVM and ANN classifiers

Overall, the results from CSP-LBP-LDA are better than those obtained using others approaches. In other words, when using CSP for feature extraction, using logarithmic band power instead of variance is more suitable. Moreover, LDA classifier is more suitable than other classifiers in the feature extracted using CSP. In future work, the proposed method will be validated using larger dataset. We consider also to perform online training and testing.

Acknowledgment

The authors gratefully acknowledge the support from the King Saud University, Riyadh, Saudi Arabia.

REFERENCES

- [1] J.P. Donoghue, "Connecting cortex to machines: recent advances in brain interfaces". *Nat Neurosci.* 5 (Suppl), 1085–1088, November 2002.
- [2] Shih JJ, Krusienski DJ, Wolpaw JR. "Brain-Computer Interfaces in Medicine". *Mayo Clinic Proceedings.* 2012;87(3):268-279.
- [3] Anderson R.A., Musallam S., Pesaran B., "Selecting the signals for a brain-machine interface", *Curr Opin Neurobiol* Vol.14 (6), pp.720-726, December 2004.
- [4] Saeid Sanei and J.A. Chambers, "EEG signal processing". Centre of Digital Signal Processing Cardiff University, UK, 2007.
- [5] Wolpaw J. R., Birbaumer N., McFarland D. J., Pfurtscheller G., and Vaughan T. M. "Brain-computer interfaces for communication and control", *Clinical Neurophysiology*, 2002, 113: 767-791.
- [6] Serby H., Yom-Tov E., and Inbar G. F., "An improved P300-based brain-computer interface", *IEEE Trans. Neural Syst. Rehabil. Eng.* 13, 89–98.
- [7] Middendorf M., McMillan G., Calhoun G., and Jones K. S. (2000), "Brain-computer interfaces based on the steady-state visual-evoked response", *IEEE Trans. Rehabil. Eng.* 8, 211–214.
- [8] Wang Y., Wang R., Gao X., Hong B., and Gao S. "A practical VEP-based brain-computer interface". *IEEE Trans. Neural Syst. Rehabil. Eng.* 14, 234–239 10.1109/tnsre.2006.875576.
- [9] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov. 1999.
- [10] Wolpaw J.R. and McFarland D.J., "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans", *Proc. Natl. Acad. Sci.* 2004;101:17849–17854.
- [11] H. Ramoser, J. M. Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehab. Eng.*, vol. 8, no.4, pp.441-446. 2000.

- [12] A. G. Baziyaad and R. Djemal, "A study and performance analysis of three paradigms of wavelet coefficients combinations in three-Class motor imagery based BCI," 2014 5th International Conference on Intelligent Systems, Modelling and Simulation, Langkawi, 2014, pp. 201-205.
- [13] R. Chatterjee and T. Bandyopadhyay, "EEG based motor imagery classification using SVM and MLP," 2nd International Conference on Computational Intelligence and Networks (CINE), At Bhubaneswar, Volume: 2, 10.1109/CINE.2016.22.
- [14] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-Robert Müller, "Optimizing spatial filters for robust EEG single-trial analysis" *IEEE Signal Processing Magazine* .
- [15] B. Blankertz et al., "The BCI competition III: validating alternative approaches to actual BCI problems," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153-159, June 2006.
- [16] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, G. Pfurtscheller, *BCI Competition 2008-Graz data set A*, Graz University of Technology, 2008.
- [17] X. Perrin, "Semi-autonomous navigation of an assistive robot using low throughput interfaces," Ph.D. dissertation, ETHZ, Zurich, Switzerland, 2009.
- [18] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification: John Wiley & Sons", 2012.
- [19] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, pp. 121-167, 1998.
- [20] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in *Encyclopedia of database system*, ed: Springer, 2009, pp. 532-538.