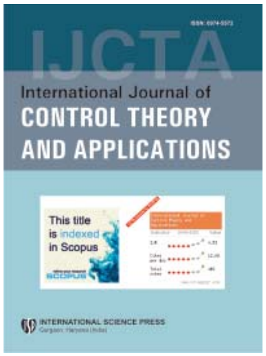


International Journal of Control Theory and Applications

ISSN : 0974–5572

Volume 9 • Number 42 • 2016

Comparison of Machine Learning Methods for Devanagari Script-based P300 Speller System

Rahul Kumar Chaurasiya[a] Narendra D. Londhe[b] and Subhojit Ghosh[b]

[a]*Department of Electronics and Telecommunication, National Institute of Technology, Raipur, India*

E-mail: rkchaurasiya@nitrr.ac.in (Corresponding author)

[b]*Department of Electrical Engineering, National Institute of Technology, Raipur, India*

E-mail: nlondhe.ele@nitrr.ac.in, sghosh.ele@nitrr.ac.in

Abstract: A P300 speller-based brain-computer interface (BCI) can connect the brain to a computer by a virtual keyboard and avoids the need for peripheral nerve and muscle activities. The communication is based on the detection of P300 event related potentials (ERPs) in the recorded electroencephalogram (EEG) signals. This paper evaluates the performances of different machine learning methods based on the classification accuracies. The EEG data were acquired from 10 subjects by using Devanagari script (DS)-based P300 speller. In primary stage, the data were preprocessed and optimum EEG channels were detected to extract the most discriminant features. Support vector machine (SVM), Fisher's linear discriminant analysis (FLDA), step-wise LDA (SWLDA), *k*-nearest neighbour (KNN) classifier and artificial neural network (ANN) were used to classify target and non-target stimuli. A multiple statistical comparison analysis carried out in the paper reveals that SVM was the overall best performing method and KNN produced the least accurate results. However, there was no significant difference between the performance of SVM, SWLDA, FLDA, and ANN.

Keywords: *P300 spellers, Brain-Computer Interface, SVM, SWLDA, KNN, Devanagari.*

1. INTRODUCTION

A brain-computer interface (BCI) is a system that connects the brain to a computer directly and avoids the need for peripheral nerve and muscle activities to execute user's actions [1]. A major aim of BCI research is to allow patients with severe motor disabilities to regain autonomy and communication abilities. A P300 speller system aims to develop a direct communication from a user to a computer machine by imitating a computer-keyboard [1-5]. The system is based on the generation and detection of P300 event related potential (ERP) in oddball paradigm based experiment [6, 7].

The first P300 speller was developed to spell English words using a 6×6 matrix of alphanumeric characters [8]. All rows and columns of the matrix were intensified in a random fashion and the subject's task was to silently count the number of times the target character was intensified. This P300 speller is commonly known as RC (row/column) paradigm-based speller [9].

Several review articles published on P300 spellers have described about the current progress and future paths [10-12]. Comparisons on various machine learning techniques for English alphabetic script-based P300 speller have also been published. [13-18].

In our previous works on P300 spellers, we proposed and developed a novel Devanagari script (DS)-based speller system [19, 20]. An 8 × 8 display paradigm consisting of 50 DS characters, 10 DS numbers and 4 special characters, was used for stimulation the subjects. The proposed paradigm can be used to input text from *Hindi, Marathi, Sanskrit, Nepali, Pali, Konkani, Bodo, Sindhi* and *Maithili* etc. languages. The EEG responses were acquired from 10 healthy subjects using a 64 channel EEG acquisition device. The main goal of the studies was to validate the fact that DS-based speller can be used for P300 based-BCI. The optimal channels were selected using binary differential evolution (BDE) algorithm. Support vector machine (SVM) classifier and its ensemble version were used to classify the target verses non-target stimuli.

This paper evaluates the performances of different other machine learning methods for the same dataset. Along with SVM, Fisher's linear discriminant analysis (FLDA), step-wise LDA (SWLDA), k-nearest neighbor (KNN) classifier, and artificial neural networks (ANN) machine learning methods have also been applied for classification.

The paper is organized as follows: section 2 describes about the EEG dataset acquired using the DS-based paradigm. Signal preprocessing, feature extraction, and optimal channel selection procedure is also described in section 2. Section 3 describes about the machine learning methodologies used for classification. Section 4 is dedicated to results and discussion. The paper is concluded in section 5.

2. MATERIALS AND METHODS

2.1. Dataset

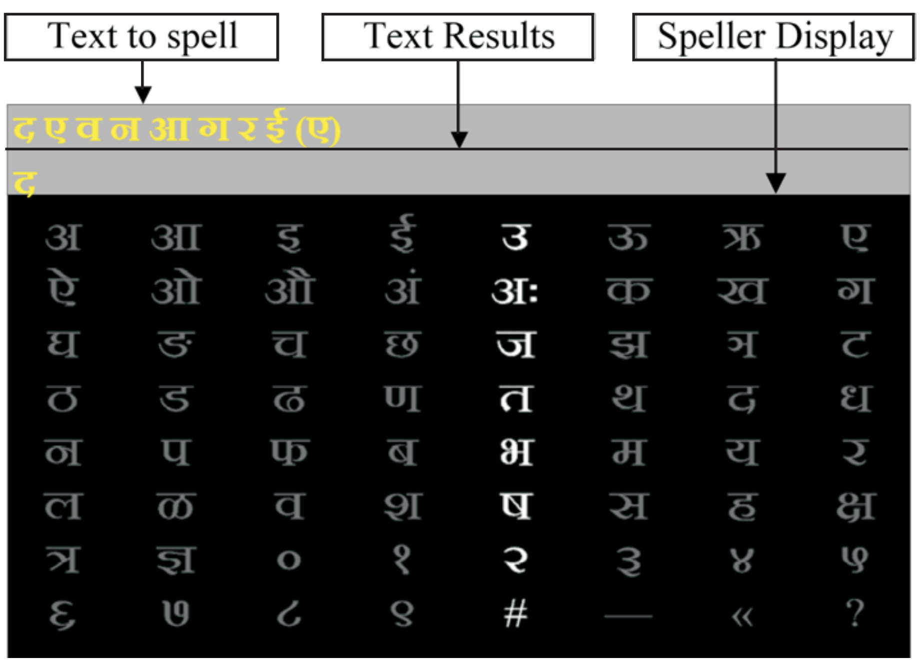


Figure 1: DS-based display paradigm for P300 speller. The paradigm contains 50 characters and 10 digits of DS and 4 special characters

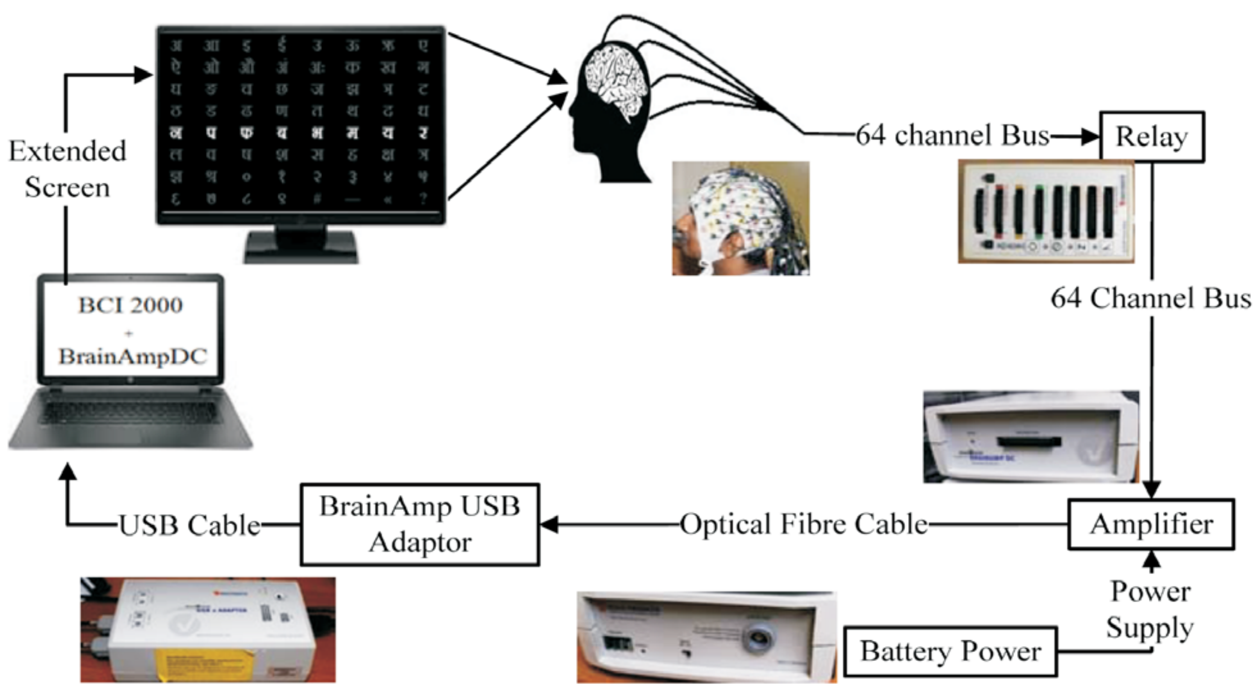
The EEG responses were recorded with 10 healthy subjects with a mean age of 26.5 (std = 2.46, range = 21-29) at Primate Research Lab, Indian Institute of Science (IISc), Bangalore, India. The subjects were not having any previous experience with BCI systems and were able to communicate using DS. The display paradigm used for DS-based P300 speller is shown in fig. 1. A 64-channel ActiCAP BrainAmp DC (of Brain Vision, UK) equipment was used to acquire the EEG response. The EEG recording setup and the electrode positions are depicted in Fig. 2. The data samples were collected with the sampling rate of 500 Hz and were digitally filtered between 1 to 250 Hz during data collection. The detailed description of the data set can be seen in [19].

2.2. Signal preprocessing and feature extraction

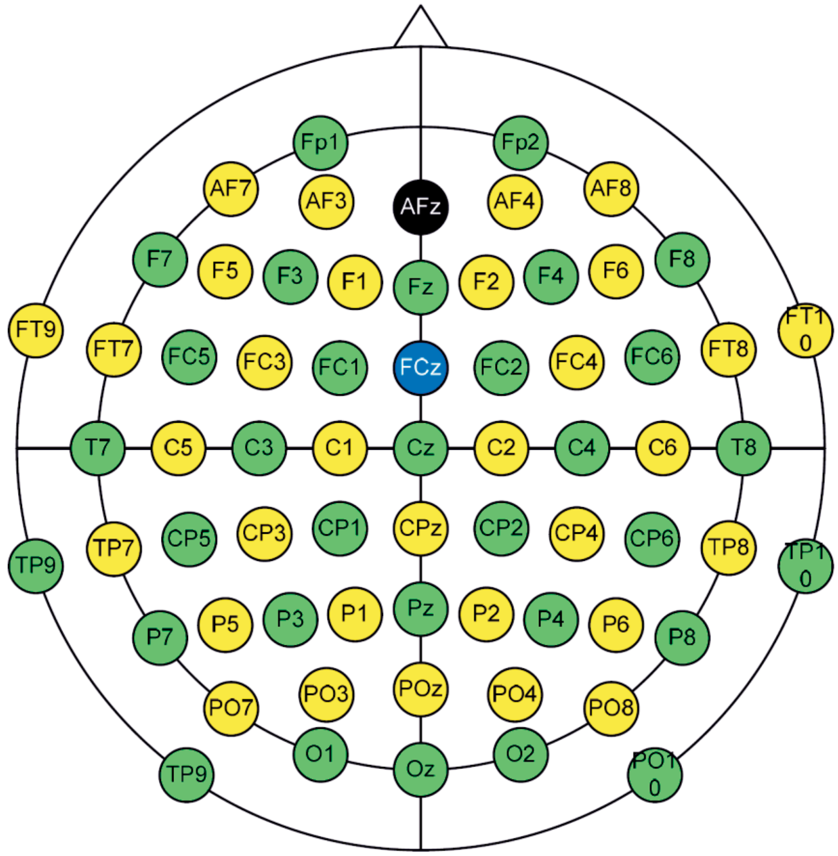
EEG signals from 0 to 600 ms posterior to each intensification of RC were extracted from each channel. The signals were then band-pass filtered with cut-off frequencies of 1-10 Hz and were decimated with frequency of 10 Hz. After preprocessing, the signals were having total 6 samples per flashing per channel. Total 15 trials were recorded for each character and hence after preprocessing, for each character, we get 240 feature vectors (16 RC × 15 trials) each of dimension 384 (6 samples × 64 channels). In each of 16 feature vectors from recorded from one trial, 2 (one row and one column) contain P300 ERP and are labelled as class +1, rest of the vectors are in class -1.

2.3. Optimized channel selection

Though the EEG responses are acquired from 64 channels from the scalp of different subjects, the data acquired from some channels might be redundant and less discriminative for P300 ERPs. Hence, for selecting the channel subset that can maximize the classification accuracy, channel selection method has been applied. The problem of selecting the best channel configuration has been formulated as an optimization problem. The problem is then solved using binary DE algorithm [21-23]. The detailed description about the binary DE algorithm-based channel selection mechanism can be seen in [19].



(a)



(b)

Figure 2: (*a*) The block diagram EEG recording set-up. (*b*) The 64-channel electrode configuration

3. MACHINE LEARNING METHODS FOR CLASSIFICATION

Due to the highly complex, noisy, and variable nature of non-invasively recorded EEG signals, the computer sometimes misinterprets the signals and makes a decision that does not match the user's intention. In this context, it is highly relevant to judge the performance of different machine learning methods for accurate detection of target characters. This section describes about the five machine learning algorithms (*i.e*. SVM, FLDA, SWLDA, KNN and ANN) applied in this paper, for classification.

3.1. SVM

SVM is one of the most widely used classifier. Being a regularized classifier, SVM can accommodate noise and outliers and increase its generalization capabilities [24, 25]. The main idea of SVM is to construct a hyper-plane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. More precisely, it is an approximate implementation of the method of structural risk minimization. SVM learns a hyper-plane

$$w^{\mathrm{T}} x + b = 0 \tag{1}$$

Where, w is the weight vector and is the bias term learned by SVM in such a way that it maximizes the separation *margin* between the two parallel hyper-planes (support vectors) described as (see figure 3).

$$w^{\mathrm{T}} x + b = 1,$$

and

$$w^{\mathrm{T}} x + b = -1 \tag{2}$$

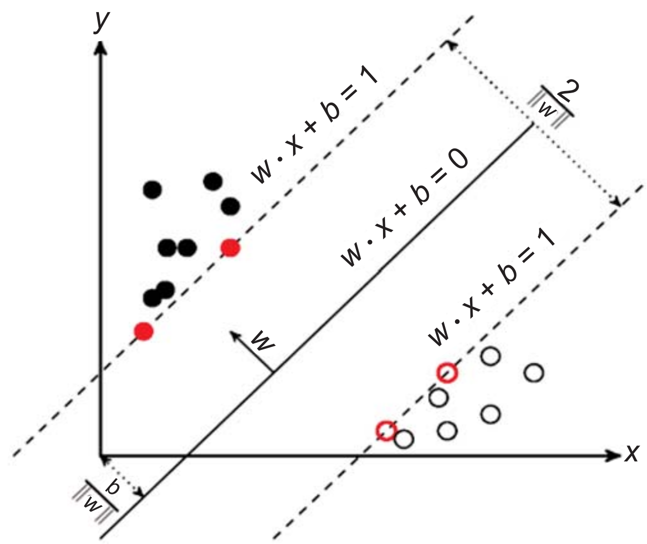


Figure 3: SVM learns a hyper-plane to maximize the separation margin between two separable classes

SVM can also act as a non-linear classifier by transforming the linear decision boundaries to nonlinear decision boundaries using "kernel trick". In the present work, a redial basis function (RBF) kernel-based SVM has been applied for classification of EEG signals.

3.2. FLDA

The FLDA (or simply LDA) [26] method is a binary classification method. Similar to linear SVM, it uses a hyper-plane to separate the data into one of the two classes. In FLDA, the data from both the classes is assumed to be normally distributed with equal covariance matrix for both the classes. To assign a class to a new feature vector , it learns a separating hyper-plane, which is a linear discriminant function that maximizes the distance between the means of the two classes and minimizes the within-class variance. The low computational complexity and simplicity of the LDA classifier allows it to produce good results at a significantly faster rate. FLDA.

3.3 SWLDA

Stepwise LDA [27] is an extension of FLDA, which uses lesser number of features for classification than FLDA. Only those features, which are more suitable for classification, are selected for discriminant analysis in SWLDA. In other words, SWLDA removes those terms from linear discriminant model, which are less significant for regression. The most significant features are selected by a combination of forward and backward stepwise regression. In the presented work, the final discriminant function was restricted to contain a maximum of 60 features.

3.4. KNN

In KNN, the assignment of the class-label to a test data point is based on the class of its nearest neighbour(s). Given a new test data point , we compute its *proximity* with each data pint in the training data set. The *proximity* here is a similarity measure between the two data pints and generally is a Euclidian distance between them. The class label of test data point is then assigned based on the majority of the class of *k*-nearest data points. KNN classifier is sensitivity to the curse-of-dimensionality.

In the present work, based on different trial and error experiments, the value of k was selected to be 10.

3.5. ANN

ANN are complex computational systems whose structure is inspired from the neural network of human brain. These systems work on the principle that a complex task can be performed by a group several neurons, similar to that of the brain. The Multilayer Perceptron (MPL) or the feed forward NN is the most famous neural network and the same has been used in this article. An MLP consists of an input and an output layers with several hidden layers between them. The output of each neuron in a particular layer is feed-forwarded as the input to the neurons of the next layer.

A number of pilot runs were carried out and architecture having two hidden layers (consists of 256 neurons and 16 neurons, respectively) was used in the present work.

3.6. Detection of target character in the test data

After the learning stage of different machine learning methods, they were applied to classify test feature vectors. Ideally, only 2 features vectors must be assigned with class +1 and 14 with class -1 (for each trial) as a result of classification. However, due to complex, noisy, and variable nature of the signals, the correct symbol may not be detected from a single trial. So a multi-trial approach was applied. After a number of trials, one out of 8 rows and one out of 8 columns (with the highest score for being in class +1 in all rows and columns, respectively) was classified to be in class +1. The target character is then decided to be the cross-section point of detected row and column (as in fig. 1).

4. RESULTS AND DISCUSSION

This section presents the results classification accuracies obtained for character detection by different machine learning methods. As suggested in [19], the EEG channels Pz, CP6, PO4, P2, CPz, PO10, Cz and PO7 are able to provide the most discriminative features across all 10 subjects. These channels are the most frequently selected channels using binary DE-based optimal channel selection method. Hence, in the present work, the data collected from these 8 channels were used to train and test the classifiers. Table 1 presents the accuracy of character detection obtained by different machine learning methods. All 15 trial data were used to train and test the classifiers. A 5-fold cross validation was [28] carried out for classification of 100 characters.

Table 1
The Accuracy character detection by different machine learning Algorithms using 15 Trials and for the dataset collected from 8 channels

Subject No.	*Accuracy*				
	SVM	*FLDA*	*SWLDA*	*KNN*	*ANN*
1.	87	83	86	75	84
2.	85	81	83	72	74
3.	88	84	85	62	80
4.	87	82	84	68	82
5.	86	88	81	74	76
6.	92	86	92	81	88
7.	82	76	84	76	87
8.	88	83	86	77	85
9.	84	85	88	80	86
10.	89	88	90	84	84
Average	86.8	83.6	85.9	74.8	82.6

4.1. Statistical Analysis

In order compare the performance of different machine learning algorithms, a multiple-comparison analysis test was applied on the classification results. For this purpose, Friedman test was employed. The test is a non-parametric equivalent of repeated-measure ANOVA and does not assume a normally distribution in the sample values [29]. Under Friedman test, the different machine learning methods are ranked according to the respective accuracy values for each subject, separately. Based on the average ranks, a Friedman test statistics is obtained. Based on the experimental results, the computed *p-value* (1.95×10^{-5}) was less than 0.01 indicating a significant difference in the performances of different approaches.

Further, a post-hoc Nemenyi test is applied to report the significant difference between the individual approaches [30]. A critical difference (CD) value of 1.929 is obtained and the performances of two methods is said to be significantly different if their rank difference is more than CD. The results of the post-hoc test are shown in figure 4. The average rank for each method is in ascending order. Horizontal lines of different colors indicate no significant difference between the methods connected by them.

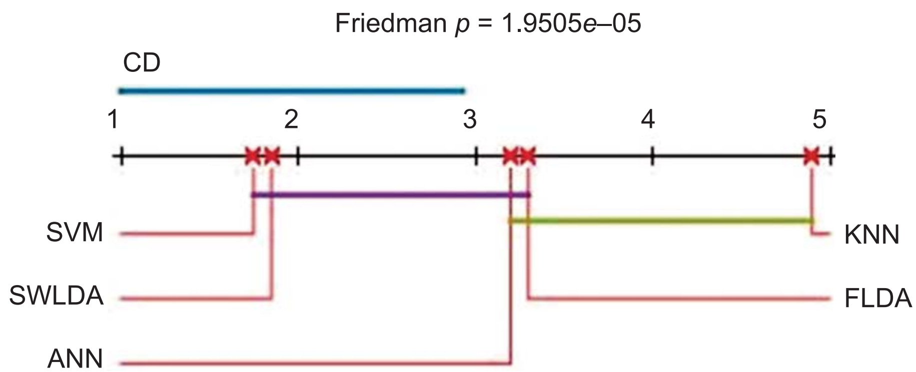


Figure 3: Visualization of Friedman test and post-hoc test for multiple comparison analysis of different machine learning methods

It can be observed from table 1 and figure 4 that on an average, SVM performed better than all other methods. However, there was no significant difference between the performance of SVM, SWLDA, FLDA and ANN methods. The worst performance was obtained by KNN algorithm. Moreover, the performance of SVM and SWLDA was significantly better than KNN.

5. CONCLUSION

The aim of this research work was to evaluate the performance of different machine learning methods for DS-based P300 speller system. Five different machine learning methods were applied to detect P300 ERP for character prediction on the data set collected from 10 subjects. SVM performed best, followed by SWLDA, FLDA, ANN, and KNN, in decreasing order of performance. A statistical significance test was also carried out to see if there is any significant difference between the performances of different methods. Though there was no significant difference between SVM, SWLDA, ANN and FLDA, the performance of SVM and SWLDA was significantly better than KNN.

6. ACKNOWLEDGEMENT

Authors acknowledge Department of Science and Technology, Government of India for financial support vide Reference No. SR/CSRI/38/2015 (G) under Cognitive Science Research Initiative (CSRI) to carry out this work.

REFERENCES

[1] J. R. Wolpaw, *et al.*, "Brain-computer interfaces for communication and control," *Clin Neurophysiol*, vol. 113, pp. 767-791, Jun 2002.

[2] J. R. Wolpaw, *et al.*, "Brain-computer interface research at the Wadsworth Center," *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, pp. 222-226, 2000.

[3] E. A. Curran and M. J. Stokes, "Learning to control brain activity: a review of the production and control of EEG components for driving brain-computer interface (BCI) systems," *Brain Cogn*, vol. 51, pp. 326-336, Apr 2003.

[4] T. M. Vaughan, *et al.*, "Brain-computer interface technology: a review of the Second International Meeting," *IEEE Trans Neural Syst Rehabil Eng*, vol. 11, pp. 94-109, Jun 2003.

[5] N. Birbaumer, "Brain-computer-interface research: coming of age," *Clin Neurophysiol*, vol. 117, pp. 479-483, Mar 2006.

[6] S. Sutton, *et al.*, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, pp. 1187-1188, Nov 1965.

[7] E. Donchin, "Presidential address, 1980. Surprise!...Surprise?," *Psychophysiology*, vol. 18, pp. 493-513, Sep 1981.

[8] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr Clin Neurophysiol*, vol. 70, pp. 510-523, Dec 1988.

[9] R. Fazel-Rezai, *et al.*, "P300 brain computer interface: current challenges and emerging trends," *Front Neuroeng*, vol. 5, p. 14, 2012.

[10] J. R. Wolpaw, *et al.*, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans Rehabil Eng*, vol. 8, pp. 164-73, Jun 2000.

[11] F. Lotte, *et al.*, "A review of classification algorithms for EEG-based brain-computer interfaces," *J Neural Eng*, vol. 4, pp. R1-R13, Jun 2007.

[12] M. Akcakaya, *et al.*, "Noninvasive brain-computer interfaces for augmentative and alternative communication," *IEEE Rev Biomed Eng*, vol. 7, pp. 31-49, 2014.

[13] D. J. Krusienski, *et al.*, "A comparison of classification techniques for the P300 Speller," *J Neural Eng*, vol. 3, pp. 299-305, Dec 2006.

[14] H. Mirghasemi, *et al.*, "Analysis of P300 classifiers in brain computer interface speller," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, 2006, pp. 6205-6208.

[15] A. E. Selim, *et al.*, "Machine learning methodologies in P300 speller Brain-Computer Interface systems," in *Radio Science Conference, 2009. NRSC 2009. National*, 2009, pp. 1-9.

[16] C. S. Syan and R. E. Harnarinesingh, "Comparison of pre-processing and classification techniques for single-trial and multi-trial P300-based brain computer interfaces," *American Journal of Applied Sciences*, vol. 7, p. 1219, 2010.

[17] F. Aloise, *et al.*, "A comparison of classification techniques for a gaze-independent P300-based brain-computer interface," *J Neural Eng*, vol. 9, p. 045012, Aug 2012.

[18] N. V. Manyakov, *et al.*, "Comparison of classification methods for P300 brain-computer interface on disabled subjects," *Comput Intell Neurosci*, vol. 2011, p. 519868, 2011.

[19] R. K. Chaurasiya, *et al.*, "Binary DE-Based Channel Selection and Weighted Ensemble of SVM Classification for Novel Brain–Computer Interface Using Devanagari Script-Based P300 Speller Paradigm," *International Journal of Human–Computer Interaction*, vol. 32, pp. 861-877, 2016.

[20] R. K. Chaurasiya, *et al.*, "A Novel Weighted Edit Distance-Based Spelling Correction Approach for Improving the Reliability of Devanagari Script-Based P300 Speller System," *IEEE Access*, vol. 4, pp. 8184-8198, 2016.

[21] R. Storn and K. Price, *Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces* vol. 3: ICSI Berkeley, 1995.

[22] L. Wang, *et al.*, "A novel modified binary differential evolution algorithm and its applications," *Neurocomputing*, vol. 98, pp. 55-75, 2012.

[23] Y. Chen, *et al.*, "A binary differential evolution algorithm learning from explored solutions," *Neurocomputing*, vol. 149, pp. 1038-1047, 2015.

[24] P. Sastry, "An introduction to support vector machines," *Computing and information sciences: Recent trends*, 2003.

[25] S. Theodoridis and K. Koutroumbas, *Pattern Recognition, Fourth Edition*: Academic Press, 2008.

[26] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, pp. 179-188, 1936.

[27] N. R. Draper and H. Smith, *Applied regression analysis*: John Wiley & Sons, 2014.

[28] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, 1995, pp. 1137-1145.

[29] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *The Journal of Machine Learning Research*, vol. 7, pp. 1-30, 2006.

[30] P. Nemenyi, "Distribution-free multiple comparisons," in *Biometrics*, 1962, p. 263.