An Efficient and High Accuracy P300 Detection for Brain Computer Interface System Based on Kernel Principal Component Analysis

Mohammad Farukh Hashmi*, Jagdish D. Kene², Deepali M. Kotambkar³, Praveen Matte⁴ and Avinash G. Keskar ⁵

¹-⁵Electronics and Communication Engineering Department,
¹National Institute of Technology, Warangal, India
²-³Shri Ramdeobaba College of Engineering and Management, Nagpur, India
⁴JSPM’S RSCOE, Polytechnic Tathawade, Pune, India
⁵Visvesvaraya National Institute of Technology, Nagpur, India
mdfarukh@nitw.ac.in, kenejd@rknec.edu, shelkedt@rknec.edu, mattepravin@gmail.com, agkeskar@ece.vnit.ac.in
*Corresponding Author: Mohammad Farukh Hashmi

Abstract: Human machine interaction with the use of brain signals has been made possible by the advent of the technology popularly known as brain computer interface (BCI). P300 is one such brain signal which is used in many BCI systems. The problems associated with most of the existing P300 detection methods are that they are time consuming and computationally complex as they follow the procedure of averaging the values obtained from multiple trials. Also the existing single trial methods have been able to obtain only moderate accuracy levels. In this paper, a novel approach which for achieving a high level of accuracy has been proposed for single trial P300 signal detection amidst noise and artifacts. In this method features were obtained by applying Discrete Wavelet Transform followed by a technique making use of the obtained wavelet coefficients. Kernel Principal Component Analysis (KPCA) was used for reducing the feature dimension. Classification of the P300 signal using the reduced features was done using Support Vector Machine (SVM). The Dataset used was the Dataset II of the third BCI Competition. An accuracy of 98.53% was achieved for Subject S1 (signal obtained from the first person) and 99.25% for Subject S2 (signal obtained from the second person) by using the proposed method. A high level of accuracy was obtained, as compared to many existing techniques. Also the speed of classification was improved with the use of reduced feature dimensions.

Keywords: Single trial P300 Detection, Brain Computer Interface Technology (BCI), Kernel Principal Component Analysis (KPCA), Discrete Wavelet Transform (DWT), Support Vector Machine (SVM).
1. Introduction

With the advancements in the field of computer technology and neuroscience, it is now possible to link brain signals to operate various applications via computers or devices. Brain Computer Interface Technology (BCI) is one of the most researched areas in the field of Biomedical Engineering. From the field of entertainment to the world of automations, BCI can play a major role to ease life of different categories of people. Various gaming applications, autonomous vehicles, devices that can help differently abled people, etc. are in the development stage. Based on the brain signal recording methods, BCI systems can be classified into two categories: invasive system and non-invasive system. In the invasive system, BCI devices are implanted directly into the brain. Though they provide the best performance in terms of accuracy, the procedure involved in its implantation is risky as well as expensive. Whereas non-invasive BCI devices are safer and less expensive but they provide poor signal to noise ratio due to the attenuation of the signal by the skull. Thus this paper primarily focuses on non-invasive systems.

In BCI systems (both invasive and non-invasive), Electroencephalography (EEG) has been found as a primary sensing technique. Among the various types of EEG signals, event-related potential (ERP) signals are mainly utilized in the design of BCI applications. ERP signals are the direct result of specific sensitive, cognitive or motor events. P300 is one such ERP signal which has been extensively used in the research. P300 is an ERP which occurs approximately 300ms after a rarely presented event or stimuli. P300 signal is one of the most popularly used signal in the design of several BCI applications such as BCI speller, Brain finger print, Lie detector, home automation systems etc. [1], [2], [3]. Due to the high importance of BCI systems, development of accurate and robust detection algorithms of the P300 signals is also very important. Although a few algorithms have already been developed, efficiency and robustness is still lacking. This paper is motivated by the fact that for proper use of BCI technology, there is a requirement of efficient and robust algorithms which can lead to more user friendly system. This was the motivation for coming up with a novel P300 detection method. The average value of a large number of trials is used in most of the existing BCI systems considering the removal of various artefacts. This process increases the signal detection accuracy but at the cost of communication rate as averaging consumes a lot of time. Most of the BCI systems use large number of electrodes for acquisition of signal which make the system costly. This is the other limitation of the existing BCI system. Yet another challenge is the fact that most of the biomedical signals are highly subject-dependent. So the characteristics of the signal vary from one individual to another.

2. Literature Review

Extensive research has been carried out in the area of P300 signal detection for real time applications. Gerardo et al. proposed a method making use of EMD (Empirical Mode Decomposition) and SVM for P300 signal detection. The method uses 9 channels for acquisition of brain signals with an achieved accuracy of 53.4% [4]. Jin-he et al. developed a P300 detection method where dynamic features are
extracted using Fisher distance and Wavelet transform (WT) classified using neural networks. 85% accuracy is obtained in this method [5]. Neda et al. suggested a detection method based on improved Independent Component Analysis (ICA), Wavelet Transform (WT) for feature extraction and Linear Discriminant Analysis (LDA) for feature reduction and classification. An average accuracy of 65% was achieved in P300 detection [6]. Yupeng et al. Suggested a P300 signal method using ICA, Fisher Distance for dynamic channel selection and Wavelet Transform. Average accuracy achieved by this method is 89.5% [7]. Ravi et al. proposed an algorithm which uses LS-SVM (Least Square Support Vector Machine). Less number of channels (i.e. five) were used and managed to obtain an accuracy of 71.5% [8]. Cecotti et al. came up with a convolutional neural network (CNN) based P300 detection method. They have proposed four single classifiers and three multi classifiers based on CNN. An accuracy of 95.5% was achieved by the method [9]. Vahid et al. proposed a two stage p300 detection algorithm. The first stage consists of deionizing of raw data using ICA-wavelet based method. The features were extracted using ICA based feature extraction method in the second stage. They used neural network for classification of the signal [10]. The accuracy obtained was 71.5%. The winners of BCI Competition III [11] who have used the same dataset managed an accuracy of 96.5% with 15 trials and 73.5% accuracy using 5 trials. The dataset used by all the above authors is taken from BCI Competition III Dataset II.

Many researchers have made studies on P300 detection by making use of the data acquired by their own experimental set up. Ozair et al. used constrained ICA with linear classifier for P300 detection. Here they took the average of many trials to obtain good results [12]. Riley et al. came up with a single trial method which used Genetic algorithm and linear discriminant analysis classifier [13]. Songyun et al. proposed method for P300 detection using extreme learning algorithm [14]. C.M. Held et al. suggested single trial P300 detection in children method which used six wave features. They used Self-Organizing Maps (SOM) based classifier [15]. Kaper et al. suggested a multi-trial method with the use of ten electrodes. They used SVM classifier to classify the preprocessed raw data [16]. Mautner et al. proposed another single trial P300 detection algorithm which used windowed mean for feature extraction. This method captures slow-changing cortical potentials. LDA was used as the classifier in this method [17]. Motlagh et al. used CWT combined with statistical T-Student (T-CWT) along with DWT for feature extraction. They also used. It was followed by SVM classification for P300 detection [18]. Lazar et al. introduced another P300 detection method which is based on morphological operators like opening, closing and weighted open-closing [19]. Turnip et al. came up with a single trial P300 detection method. Features are extracted using adaptive non-linear principal component analysis and classified using neural networks [20]. Achieving higher accuracy with speed is the prime target of current BCI researches. Reaching above goals with single trial detection is one of the most popular areas of research today. The proposed method uses single trial unlike the various existing methods which are using multiple trials, to reduce the computation time to a large extent. Also, use of six electrodes in this approach makes it more cost-effective and user friendly to the user. In this method features are obtained by applying Wavelet Transform followed by a new technique in which the obtained wavelet coefficients are divided into certain number of sets and among them certain specific coefficients are chosen and used as the feature vector. Subsequently, a feature dimension reduction technique KPCA is applied on the obtained features in order to further reduce the computation complexity. SVM is used to classify the signal using the reduced features. The set of performance parameters such as sensitivity, specificity, accuracy and precision were used to evaluate to the performance of the proposed method.
3. Materials

3.1 Dataset

The dataset supplied by BCI Competition III data set II for P300 speller paradigm is used in this experiment [21]. This dataset consists of training samples as well as testing samples. Among these we have used only the training samples for our study. This is a standard benchmark dataset which has been used in developing many of the P300 detection algorithms. In order to compare the performance of our developed algorithm with the existing ones we have used the same dataset provided by BCI Competition III. This dataset consists of data taken from only two subjects during an experiment. In this experiment a 6x6 matrix (Fig.1) was presented in front of each subject which contains 36 symbols. The matrix consists of 26 alphabets and other useful characters. The user was made to wear headset consisting of 64 electrodes. The user was asked to concentrate on a particular symbol which he wants to communicate. The 6x6 symbol matrix consists of 6 rows and 6 columns. The backgrounds of rows and columns were intensified one at a time in a random manner for a fixed period of 100ms. A blank matrix is presented for duration of 75ms after each intensification attempt. The cycle of 12 intensifications covering each row and column once is called a sequence which was repeated for 15 times. The Subject stares at one of the 36 characters in the matrix at a given time and maintains a mental count. We can notice that in a sequence, only two intensifications (corresponding row and column containing the symbol) were the target intensifications. Hence the target character was intensified twice a sequence and 30 times in 15 sequences. 85 symbols or characters were targeted by each subject. The same process of 15 sequences was repeated for all the targeted symbols or characters. An ERP, P300 potential is generated whenever the targeted row or column (containing the targeted symbol) was intensified. The 64 electrodes cap which was placed over the subject's head will measure the ERP at a sampling rate of 240 Hz [6] as shown in Fig.2. We have used the data from only six channels (Fz, Pz, P1, P2, POz and Oz) out of the 64 channels. The P300 signals are strongly generated in the specific area of scalp which is parental lobe [21]. The above mentioned six channels read data from same area. Hence the electrode position surrounding the area mentioned in [21] having stronger P300 signals was selected based on our experimental observation.
4. Proposed Methodology

The flow of the algorithm is as shown in Figure 3. Most of the existing works were using averaged data from large number of trials in order to avoid noise and intensify P300 signal feature present in the data thereby resulting in improved accuracy. But the consequence of averaging is large time consumption which will result in decreasing the speed of the system. Hence we propose a P300 signal detection
algorithm based on single trial. KPCA and SVM are used together in our method which is uncommon in P300 detection. The proposed method uses only six electrodes for signal acquisition and processing thereby proving to be cost effective.

Signal X(n) is acquired (in the duration of 0 ms to 666.67 ms) after stimuli (from all 6 channels)

DWT is applied

Coefficients are divided into 5 parts and maxima and minima from each are taken (Feature Vector)

Corresponding feature vector from all the 6 channels are concatenated

KPCA is used for reduction of Feature dimension

SVM is used for Classification

Figure 3. Proposed Algorithm Flow

4.1 Feature Extraction

One of the most suitable methods for feature extraction of P300 signal is wavelet transformation. The Wavelet transform having multi resolution property contains frequency as well as time domain information. As P300 signals are the non-stationary and random in nature, we have selected wavelet transform as the technique to extract features of the signal. Wavelet coefficients of each data samples are obtained from the wavelet decomposed signal.

\[
D_{j,k}(i) = \sum_{j=0}^{L-K} g(k)A_j(2i + k) \tag{2}
\]

\[
A_{j,k}(i) = \sum_{j=0}^{L-K} h(k)A_j(2i + k) \tag{3}
\]

Where D(i) is the detail coefficient and A(i) is the approximation coefficient. Among these detail coefficients (D) were used for feature vector. The reason behind this is that detail coefficients
correspond to the sharp features which can more distinctly classify the signal as P300 and non-P300 unlike approximation coefficients.

Previous studies suggested Daubechies wavelet ‘db4’ at level 2 as the most suitable wavelet for EEG signals, so it was used in the algorithm.

A new technique for feature extraction is applied instead of using the wavelet coefficients directly. The wavelet coefficients (D) obtained from each signal data (160 samples) are divided into 5 different groups (S_i) consisting of 32 samples in each group.

\[ D = \{D_1, D_2, D_3, \ldots, D_{160}\} \]

\[ D = \{S_1, S_2, S_3, S_4, S_5\} \]

\[ S_i = \{D_{(p+1)}, D_{(p+2)}, \ldots, D_{(p+32)}\}, \text{i=1,2,...,5} \]

Where \( p = (i-1) \cdot 32 \)

A maximum value \( \text{max}_i \) and a minimum value \( \text{min}_i \) are obtained from each group. The obtained feature vector (of length 2 samples) consists of maxima value and minima value of each signal data.

\[ \text{max}_i = \text{maximum} (S_i) \]

\[ \text{min}_i = \text{minimum} (S_i) \]

The feature vectors from each of the 6 channels are concatenated to form a new feature vector \( \text{F} \) (2 \* 5 \* 6 samples).

### 4.2 Feature Reduction

Principal component analysis (PCA) is one of the most famous statistical techniques for dimensionality reduction and feature extraction. It linearly projects the data sample from a large number of correlated variable spaces into less number of uncorrelated variable spaces. The obtained uncorrelated variables are known as principal components. The largest possible variance is associated with the first principal component and the variance of the succeeding components keeps decreasing while maintaining orthogonality to the preceding components [13]. Data dimension is reduced by removing some of these components in the signal.

Most of the high dimensional datasets are observed to be non-linear in nature. In such cases PCA cannot be used to model the variability of data. To address such problems, a non-linear dimensionality reduction technique known as KPCA was designed. KPCA has all the advantages of the regular PCA. It gives better recognition rate and improved performance as compared with its linear counterpart. It provides an implicit non-linear mapping to a feature space where the features representing the structure in the data may be extracted better. KPCA performs data compression by reducing the data dimension while preserving the information. Data compression can be achieved using KPCA process as follows:

**Step 1:** A kernel mapping \( K(x, y) \) is chosen.
Step 2: Input data of dimension N is mapped into some nonlinear feature space \( \phi(x_i) \) where its mean is zero.

Step 3: The covariance matrix is obtained using following formula:

\[
C = \frac{1}{N} \sum_{i=1}^{N} \phi(x_i) \phi(x_i)^T
\]  
(10)

Step 4: The normalized kernel matrix is calculated:

\[
\tilde{K} = K - 2A_{1/N}K + A_{1/N}KA_{1/N}
\]  
(11)

Where \( A_{1/N} \) is a matrix with all elements 1/N.

Step 5: The Eigen vectors \( \alpha_i \) of the covariance matrix is calculated by solving the following equation:

\[
\tilde{K}\alpha_i = \lambda_i\alpha_i
\]  
(12)

Where \( \lambda_i \) are the Eigen values.

Step 6: The data with reduced dimension r is given by:

\[
Z_p = \sum_{j=1}^{N} \alpha_{pj}K(x,x_j), \quad p=1,\ldots,r
\]  
(13)

In this method the obtained feature dimension is very large. This is likely to cause problems for the SVM classifier as the detection accuracy decreases when the feature vector dimension nears the number of testing samples. So using feature reduction technique KPCA to the obtained feature vector improves the accuracy. Also KPCA process enhances the contrast between the two classes of signals namely P300 and non-P300 thus improving the performance of the method.

The non-linear kernel function used in the KPCA method is

\[
K(x,y) = \phi(x_i) \phi(y_j)^T
\]  
(14)

KPCA is applied to each feature vector \( F \) to form a new feature vector \( F' \) of dimension three. So the reduced dimension of dataset is six. Hence the final dimension of P300 signal dataset and non-P300 signal dataset used are 2550x6 and 2465x6 respectively.

Reduced P300 Dataset =

\[
\begin{bmatrix}
F'_{1,1} & F'_{1,2} \\
F'_{2,1} & F'_{2,2} \\
M & M \\
F'_{3,1} & F'_{3,2}
\end{bmatrix}_{2550 \times 6}
\]  
(15)

Where Z is the total number of signal data.
KPCA works only on two dimensional data. For offline processing, the available P300 and Non-P300 data are arranged in respective rows to form two dimensional data. The testing data (P300 or Non-P300) is appended on the corresponding two dimensional data (P300 or Non-P300) for applying KPCA feature reduction process. For online processing, information regarding input data is unknown. Hence two dimensional data is made out of input one dimensional data by using the equation provided in [25].

4.3 Classification of Signal
Studies show SVM as one of the most suitable classifier for P300 signal detection [26]. SVM classifier is a binary classifier which is used to classify signal as P300 or non-P300 signal in the proposed method. Different kernel functions like ‘Gaussian’, ‘rbf’, ‘poly’ etc. were applied to the non-linear SVM using the same dataset [23], out of which ‘rbf’ kernel function gave the best performance. Hence rbf’ kernel function is used in this paper. RBF kernel function is defined as [27]:

\[ k(x,x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \]  \hspace{1cm} (16)

Where \( \sigma \) is the corresponding parameter whose optimum value was obtained by performance assessment and systematic variation. The optimized value of \( \sigma \) was found to be 0.8. The default value ‘1’ was used for the regularization parameter \( c \), which gave optimum performance. Various experiments were carried out to obtain optimum values for ‘\( \sigma \)’ and ‘\( c \)’ using the same dataset [23]. The dataset consisting of 2550 P300 signal data and 2465 non-P300 signal data were divided randomly into two equal parts each. First half of P300 signal data and non-P300 signal data were used for training and the second half of P300 signal data and non-P300 signal data were used for testing. After the initial training phase of the algorithm, the classification of the testing data set given to the algorithm can be performed efficiently.

5. Results and Discussions
The dataset containing P300 and non-P300 data signals were divided randomly into two equal parts each for training and testing as mentioned above. This ensures that in each testing trial the data signals present in the training and testing dataset will be different. The average output value of the performance parameters of 10 such training and testing sets were taken as the result. The accuracy obtained was 98.62% for Subject S1 and 99.22% for Subject S2 for the proposed method. The accuracy increased with reduction in feature dimension by applying KPCA. The maximum accuracy was obtained at a reduced feature dimension of six for both the subjects. The proposed algorithm was tested with different other features. However the best results were obtained using the proposed features. The results obtained from proposed method have been compared with that of literature survey as well as different other methods which we have experimented. The above comparisons show the advantages of the proposed method over existing ones.
5.1 Signal Analysis

The processing of algorithm and various technical computing has been done using MATLAB 2018, run on Intel core i5 processor. A lot of variations in the features were experimented in order to reach the optimum solution with respect to the maximum accuracy. The time domain plots of detected P300 and non-P300 signals are as shown in Fig.4 and Fig.5 respectively. These signals were obtained using Pz electrode and were obtained from a single trial. The extracted signal from beginning of the stimuli upto 666.67 ms has been plotted in Fig.6. The positive deflection peak in the P300 signal plot is clearly distinguishable as compared to non-P300 signal plot. The peak of P300 is observed at 560.3 ms after the visual stimuli. The P300 signal usually appears anywhere between 300ms to 700 ms practically and is clearly distinguished from the normal signal by a positive peak like the one which can be observed in Fig.4.

![Figure 4. Signal Analysis for P300](image)

![Figure 5. Signal Analysis for Non-P300](image)
5.2 Changes In The Detection Accuracy of P300 Signal With Feature Dimensions

During the experiments it was observed that there was larger increase in accuracy when the feature reduction technique KPCA was applied instead of directly using the features. Table 1 shows the variation in accuracy with the changes in reduced dimension of the feature vector.

| Feature Vector Dimension | % Accuracy for Subject S1 | % Accuracy for Subject S2 |
|--------------------------|---------------------------|---------------------------|
| 14                       | 96.49                     | 90.1                      |
| 12                       | 97.46                     | 93.92                     |
| 10                       | 96.9                      | 97.94                     |
| 8                        | 98.1                      | 97.58                     |
| 6                        | 98.63                     | 99.22                     |
| 4                        | 97.37                     | 97.95                     |
| 2                        | 94.45                     | 95.85                     |

Figure 6. Changes in the Accuracy of P300 signal with change in dimension of reduced feature vector

A lot of variations in the feature vector dimension were experimented and it was observed that the maximum accuracy was achieved with the reduced dimension of feature vector, six for both the subjects S1 and S2 as shown in Fig.6. It has to be noted that the values given in Table.1 are the average values obtained from 10 successive trials, after randomizing testing and training dataset before each trial as mentioned in the previous section.

5.3 Changes in the accuracy of P300 signal Detection with Different Features

Different types of features were extracted and classified using SVM classifier. Table 2 shows the performance of different features were compared while detecting P300 signal with respect to accuracy. Wavelet features were the least effective in detecting the P300 signal while EMD features and peak specifications were comparatively better than wavelet.
Table 2. Variation of accuracy with Features used

| Features Used          | % Accuracy for Subject S1 | % Accuracy for Subject S2 |
|------------------------|---------------------------|---------------------------|
| EMD                    | 84.08                     | 86.16                     |
| Peak Specification     | 82.59                     | 79.49                     |
| Wavelet with KPCA      | 67.76                     | 68.55                     |
| Proposed Algorithm     | 98.63                     | 99.22                     |

From the table it can be observed that the wavelet coefficients features were showing average performance in detecting the P300 signal. The modified wavelet features along with KPCA and SVM Classifier have shown a reasonable increase of 31.77% in accuracy. The above observation is clearly visible in Fig.7 where the modified wavelet features used in the proposed method give clear dominance performance in comparison with the other feature vectors.

![Figure 7](image)

Figure 7. Changes in the accuracy of P300 Signal Detection with Different Features

5.4 Performance Parameter evaluation

Performance of the classifier is evaluated based on various statistical parameters like:

**Precision**: The percent measure of correct positive predictions (P300) out of total positive predictions and can be calculated as

\[
\text{Precision} = \frac{TP}{TP + FP}
\]  
(17)

**Sensitivity**: The percent measure of correct positive predictions out of positive labelled (P300) instances and is given by
Sensitivity: \[ Sensitivity = \frac{TP}{TP + FN} \] (18)

Accuracy: The percent measure of total correct predictions (including both P300 and non-P300) and can be calculated as

\[ \text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \] (19)

Specificity: The percent measure of correct negative predictions out of negative labelled instances (non-P300) and can be calculated as

\[ \text{Specificity} = \frac{TN}{TN + FP} \] (20)

Where:
- TP: True Positive is the number of P300 signals that are correctly detected.
- FP: False Positive is the number of non-P300 that is classified as P300.
- TN: True Negative is the number of non-P300 signals that are classified as non-P300.
- FN: False Negative is the number of P300 that are that are classified as non-P300.

The performance parameter evaluation of the proposed method is given in Table 3. This is obtained based on the Confusion matrix which consists of parameters like TP, FP, TN and FN.

| Method            | Subject | TP  | FP  | TN  | FN |
|-------------------|---------|-----|-----|-----|----|
| Proposed Algorithm| S1      | 1275| 35  | 1240| 0  |
|                   | S2      | 1271| 16  | 1259| 4  |

5.5 Comparison of performance parameters w.r.t. the various feature extraction techniques

Various feature extraction techniques are compared using the above mentioned four performance parameters. It can be observed from Table 4 that proposed method has given the superior performance as compared with the other feature extraction techniques. Approximately 4.18% improvement in Precision, 24.70% improvement in Sensitivity, 2.9% improvement in Specificity and 13.8% improvement in Accuracy on an average is observed for both the subjects S1 and S2 as compared with other methods.

| Method            | Subject | TP  | FP  | TN  | FN |
|-------------------|---------|-----|-----|-----|----|
|                   | S1      | 1275| 35  | 1240| 0  |
|                   | S2      | 1271| 16  | 1259| 4  |
A comparison has been made of the accuracy of the proposed method with different existing methods as shown in Table 5. All the methods considered for comparison have used the same dataset provided by BCI Competition III which is also used by the proposed method. Hence the same benchmark has been maintained for comparison of various existing methods.

Table 5. Comparison of Accuracy with already existing methods

| Features Used                    | % Accuracy for Subject S1 | % Accuracy for Subject S2 |
|----------------------------------|---------------------------|---------------------------|
| ICA & WT                         | 85                        | 94                        |
| Raw data + LS-SVM                | 70                        | 71                        |
| WT & Fisher Distance             | 85                        | 85                        |
| Improved ICA + WT               | 65                        | 65                        |
| EMD + SVM                       | 58                        | 53.9                      |
| Proposed Algorithm              | 98.63                     | 99.22                     |

A comparison of the accuracy of the proposed method along with that of some of the popular existing methods in the literature survey has been shown in Fig.8. An improvement of at least 9.45% w.r.t all other methods are observed. The results show that the suggested method outperforms the others comprehensively. The comparison of proposed method with ICA & wavelet method (best existing method) [4] is shown in Table 6. It can be seen that the proposed method shows clear superiority with respect to parameters like accuracy, number of electrodes.
Also the proposed method has made use of the same parameters for both the subjects unlike the best existing method where subject specific parameters have been used. Yet, the proposed method has managed to achieve a high accuracy. The winners of BCI Competition III who have used the same dataset managed an accuracy of 96.5% with 15 trials and 73.5% accuracy using 5 trials. The proposed method has achieved much higher accuracy of 98.92% on an average for both the subjects with only a single trial.

Biomedical signals are normally very weak signals corrupted by noise and artifacts. The proposed algorithm proves very effective since it takes into account these factors and has inculcated necessary steps to enhance the signal quality. The pre-processing step removes most of the artifacts and noise contents in the signal. The wavelet transform is renowned for its quality performance with non-stationary signals. Hence is very effective for feature extraction in case of non-stationary signal P300 signals as well. KPCA produces the significant principal components of the features which gives way for efficient classification of the signal. SVM which has been proved to one of the best classifier in BCI systems has been used for the high performance of the method.

Table 6. Comparison with the best existing method or Algorithm

| Method            | Number of Electrodes | Reduced Feature Dimension | Accuracy (%) |
|-------------------|----------------------|---------------------------|--------------|
|                   | S1                   | S2                        | S1           | S2           |            |
| ICA & wavelet     | 13                   | 12                        | 42           | 40           | 85          | 94          |
| Proposed Method   | 6                    | 6                         | 6            | 6            | 98.62       | 99.22       |
| Improvement (%)   | **53.85**            | **50.00**                 | **85.71**    | **85.00**    | **13.62**   | **5.22**    |
6. Conclusions and Future Work

A method which makes use of modified wavelet coefficients along with KPCA and non-linear SVM classifier is proposed for detection of P300 signal. The proposed method is using the signal obtained from single trial only. Hence making the algorithm computationally fast. The developed method is making use of data acquired from only six electrodes. Thus making the system cost effective as readily available headsets like emotiv epoc, neuro headset etc. with less number of electrodes are sufficient for signal acquisition unlike 64 electrode headsets which are expensive and less comfortable to wear. Thus, an efficient method which is well optimized w.r.t parameters like number of electrodes, accuracy, number of trials etc has been developed. One of the limitations is that the dataset considered here consists of data obtained from just two subjects. But the dataset used in this paper is a standard one which has been used by many of the researchers in developing P300 detection algorithms. Hence same benchmark is defined for comparison of the proposed method with all the existing methods. Due to the limitation in the available standard dataset, this paper has considered data collected from only two subjects. In future work we would like to test our algorithm on more number of subjects. Further work will also include development of efficient BCI applications based on the proposed algorithm with higher data rate and efficiency.

Declarations:

Funding: No funding
Conflicts of interest/Competing interests: Not applicable
Availability of data and material: Not applicable
Code availability: Not applicable
Authors' contributions: Not applicable

References

[1] Abootalebi, Vahid, Mohammad Hassan Moradi, and Mohammad Ali Khalilzadeh. "A new approach for EEG feature extraction in P300-based lie detection." Computer methods and programs in biomedicine 94, no. 1 (2009): 48-57.
[2] Farwell, Lawrence A. "Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials." Cognitive neurodynamics 6, no. 2 (2012): 115-154.
[3] Holzner, Clemens, Christoph Guger, Gunter Edlinger, Christoph Gronegress, and Mel Slater. "Virtual smart home controlled by thoughts." In Enabling Technologies: Infrastructures for Collaborative Enterprises, 2019. WETICE'09. 18th IEEE International Workshops on, pp. 236-239. IEEE, 2019.
[4] Teodoro Solis-Escalante, Gerardo Gabriel Gentiletti and Oscar Ya’nez-Suarez, “Single Trial P300 detection based on the Empirical Mode Decomposition”, Proceedings of the 28th IEEE EMBS Annual International Conference, New York City, USA, Aug 30-Sept 3, 2016.
[5] K. Manikanda Kumaran and M. Chinmadurai, “A Competent Ad hoc Sensor Routing Protocol for Energy Efficiency in Mobile Wireless Sensor Networks”, Wireless Personal Communications https://doi.org/10.1007/s11277-020-07741-0
[6] Neda Haghighatpanah, Rasoul Amirfattahi, Vahid Abootalebi and Behzad Nazari, “A single channel-single trial P300 detection algorithm”, Electrical Engineering (ICEE), 21st Iranian Conference, 2013, Pages: 1-5.
[7] Yupeng Wang, Jizhong Shen, Jianwei Liang and Yu Ji, “Research of P300 Feature Extraction Algorithm based on ICA and Wavelet Transform”, Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics,year: 2014, Volume: 1, Pages: 41-45.
[8] Vanitha Narayan Raju, In-Ho Ra and Ravi Sankar, “A P300-Based BCI Classification Algorithm Using Least Square Support Vector Machine”, International Journal of Software Engineering and Its Applications Volume: 9, No. 5 (2015), pages: 247-254.
Cecotti, Hubert, and Axel Gräser. "Convolutional neural networks for P300 detection with application to brain-computer interfaces." Pattern Analysis and Machine Intelligence, IEEE Transactions on 33, no. 3 (2011): 433-445.

Haghhighatanah, Neda, Rasoul Amirfattahi, Vahid Abootalebi, and Behzad Nazari. "A two stage single trial P300 detection algorithm based on independent component analysis and wavelet transforms." In Biomedical Engineering (ICBME), 2012 19th Iranian Conference of, pp. 324-329. IEEE, 2012.

http://www.bbci.de/competition/iii/results/

Haghighatpanah, Neda, Rasoul Amirfattahi, Vahid Abootalebi, and Behzad Nazari. "A two stage single trial P300 detection algorithm based on independent component analysis and wavelet transforms." In Biomedical Engineering (ICBME), 2012 19th Iranian Conference of, pp. 324-329. IEEE, 2012.

Orhan, Umut, and Emre Gurbuz. "Classifying discrete interval densities of EEG signals by using DWT and SVM." In Innovations in Intelligent Systems and Applications (INISTA), 2012 International Symposium on, pp. 1-4. IEEE, 2012.

Manikandan S, Chinnadurai M., Thiruvengadatsubesh M.P., Sivakumar M. (2020). "Prediction of Human Motion Detection in Video Surveillance Environment Using Tensor Flow", International Journal of Advanced Science and Technology, 29(05), 2791 - 2798. Vol. 29 No. 05 (2020): Vol. 29 No. 05 (2020)

Xie, Songyun, You Wu, Yunpeng Zhang, Juanli Zhang, and Chang Liu. "Single channel single trial P300 detection using extreme learning machine: Compared with BPNN and SVM." In 2014 International Joint Conference on Neural Networks (IJCNN), pp. 544-549. IEEE, 2014.

Morales, C., C. M. Held, P. A. Estevez, C. A. Perez, S. Reyes, P. Peirano, and C. Algarin. "Single trial P300 detection in children using expert knowledge and SOM." In Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, pp. 3801-3804. IEEE, 2014.

Kaper, Matthias, Peter Meinicke, Ulf Grossekatheofer, Thomas Lingner, and Helge Ritter. "BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm." Biomedical Engineering, IEEE Transactions on 51, no. 6 (2004): 1073-1076.

Motliah, F., Sai Hong Tang, and Omid Motliah. "Optimal accuracy and runtime trade-off in wavelet based single-trial P300 detection." In Signal and Image Processing Applications (ICSIAPA), 2013 IEEE International Conference on, pp. 343-348. IEEE, 2013.

Lazar, Anca Mihaela, and Radu Ursulean. "The P300 event-related potential detection-A morphological approach." In E-Health and Bioengineering Conference (EHB), 2013, pp. 1-4. IEEE, 2013.

Manikandan, S. & Chinnadurai, M 2019, ‘Intelligent and Deep Learning Approach OT Measure E-Learning Content in Online Distance Education’, The Online Journal of Distance Education and e-Learning, vol.7, issue 3, July 2019, ISSN: 2147-6454.

Colwell K.A, Ryan D.B, Throckmortan C.S, Sellers E.W and Collins L.M. “Channel selection method for p300 speller”, Journal of neuroscience methods, Vol. 232, pp. 6-15. Elsevier, 2014.

Alfauuri, Mikhled, and Khaled Daqrouq. "ECG signal denoising by wavelet transform thresholding." American Journal of applied sciences 5, no. 3 (2008): 276-281.

BCI Competition III Data set II, Wadsworth Center, and NYS Department of Health. Available: http://www.bbci.de/competition/iii/#data_set_ii

Mammone, Nadia, Fabio La Foresta, and Francesco Carlo Morabito. "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA." Sensors Journal, IEEE 12, no. 3 (2012): 533-542.

Teixeira, A. R., A. M. Tomé, E. W. Lang, R. Schachtner, and K. Stadlthanner. "On the use of KPCA to extract artifacts in one-dimensional biomedical signals." In 2006 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing, pp. 385-390. IEEE, 2006.

Tayeb, Salma, Abdelhak Mahmoudi, Fakhita Regragui, and Mohammed Majid Himmi. "Efficient detection of P300 using Kernel PCA and support vector machine." In Complex Systems (WCCS), 2014 Second World Conference on, pp. 17-22. IEEE, 2014.

Kuo, Bor-Chen, Hsin-Hua Ho, Cheng-Hsuan Li, Chih-Cheng Hung, and Jin-Shiuh Taur. "A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification." Selected Topics in applied Earth Observations and Remote Sensing, IEEE Journal of 7, no. 1 (2014): 317-326.