“Classification of EEG Signals Using Wavelet Transform and Hybrid Classifier For Parkinson’s Disease Detection”

1, Miss. Priyanka G. Bhosale, 2, Prof. Dr. S.T. Patil.
1,2, Department of Computer Science and Engineering and Information Technology,
Vishwakarma Institute of Technology,
Pune, India.

Abstract
Feature extraction and classification of Electroencephalograph (EEG) signals for normal and abnormal person is a challenge for engineers and scientists. Various signals processing techniques have already been proposed for classification of non-linear and non-stationary signals like EEG. In this work, Support Vector Machine (SVM) and Multilayer perceptron (MLP) based classifier was employed to detect Parkinson’s disease from background electroencephalograph signals. Signals where preprocessed, decomposed by using discrete wavelet transform (DWT) till 5th level of decomposition tree. The proposed classifier may show the promising classification accuracy.

Keywords- Electroencephalograph (EEG), Support Vector Machine (SVM), Multilayer perceptron (MLP), Discrete Wavelet Transform (DWT), Parkinson’s Disease (PD).

1. INTRODUCTION
The growth in the population age brings the growth of number of diseases. And one of the major group of diseases that affect elderly people is that neurodegenerative diseases. The Parkinson Disease (PD) is disorder of certain nerve cells in the part of the brain which produces dopamine. These nerve cells break down, dopamine levels drop and brain signals which are responsible for the moment become abnormal. PD usually begins in the middle or late life (after age 50). It progresses gradually for 10-15 years. This results in more and more disability. Patients suffering from PD present more clinical abnormalities of moment like resting tremor, rigidity, bradykinesia and postural instability. Electroencephalograph (EEG) is the recording of electrical activity along the scalp, produced by the firing of neurons within the brain. In clinical context EEG refers to recording of the brain’s spontaneous electrical activity over a short period of time, as recorded from multiple electrodes placed on the scalp. In this paper the main diagnostic application of EEG is in the case of Parkinson’s disease. [1] Brain patterns form wave shapes that are commonly sinusoidal; usually they are measured from peak to peak and normally range from 0.5 to 100 µV in amplitude. Brain waves has been classified into four basic groups or bands depending on the frequency range (as shown in Fig. 1). [2]
- β beta (>13 Hz),
- α, alpha (8-13 Hz),
- θ, theta (4-8 Hz),
- δ, delta (0.5-4 Hz).

Fig. 1 Brain wave samples with dominant frequencies belonging to β, α, θ, and δ bands respectively (upper to lower).
Two basic steps must be performing to analysis of EEG signals: Feature Extraction, & Signal Classification. Feature extraction can be calculated based on statically characteristics or syntax description components of domain, frequency, time and time-frequency domain, can be used to extract features from EEG signals. Wavelet transform is a powerful mathematical tool that can be used to analysis of non-stationary signals such as EEG. It has property of time-frequency resolution and can be apply to extract various features. [2] In past few years many research groups focused their work on classifying EEG records to desired mental task classes. Several algorithms has been investigated by purpose or increasing the classification rate and accuracy. In this work an algorithm based on Daubechies Discrete wavelet transform (DWT- db) and Hybrid classifier (combination of SVM & MLP) is used to detect the PD signals from EEG signals. The rest of the paper is organized as follows section II contain methodology, section III Future Work, and in section IV contain conclusion.

2. METHODOLOGY

![Fig. 2 Proposed System Framework.](image)

The steps involved in the BCI framework

1. **Acquisition of Brain Activity** :- Many kinds of electrodes like EEG cap are used to record human brain activity in this step.

2. **Analog to Digital Conversion** :- After capturing the EEG signals which are in the analog form then it can be converted into digital form by using Analog-to-Digital converter.

3. **Signal Preprocessing** :- After capturing the EEG signals which contain the noise and other artifacts such as eye blink, moment of eyeball etc. so that there need to clean them. In this step we clean the EEG signal.

4. **Feature Extraction** :- In this step particular set of signal attribute or properties are captured and after that most relevant set of selected . In this paper we discuss the Discrete Wavelet Transform method for the feature extraction from EEG signals.

5. **Classification/ Feature translation** :- In this step the most relevant set of selected features is classified to identify the mental state. There are many classifiers are present to classify the EEG signals but here we discuss the hybrid classifier (which is the combination of SVM & MLP) to classify the features.

6. **Translation into command** :- After classification we get the signal belongs to which class i.e. in the case of PD patient if the frequency is > 30Hz then that person is normal and if the frequency is below the 30Hz then the person is abnormal. If the person is abnormal the according to frequency we can check that in which stage that person is according to their frequency.
3. Wavelet Analysis And Feature Extraction

Visual analysis and diagnosis of EEG signals using time domain analysis is time consuming and tedious task. It may vary from person to person. Frequency domain analysis also have limitation like, no information about where frequencies are located in time, more samples to analyze for getting accurate results, large memory space required for storage of data, extensive processing time, more filter length, non linear phase and lack of artifact removal. So that we can use the time-frequency (wavelet) analysis[9]. The wavelet transform form a general mathematical tool for signal processing with many application in EEG data analysis. The wavelet transform itself work as artifact removal involves the breaking down of brain signals into shorter reads of band as per requirements. In Discrete wavelet Transform, a multi-resolution description is used to decompose a given signal into increasing finer details based on two set of basic function [9]. The wavelet and scaling function as follows

\[ f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_j^k \psi(2^j t - k) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d_j^k \varphi(2^j t - k) \]  

(1)

Where function \( \varphi(t) \) & \( \psi(t) \) are the basic scaling and mother wavelet respectively. In above expression the first summation present approximation \( f(t) \) based on scale index of \( j_0 \), while the second term add more detail using larger \( j \) (finer scale). The second coefficient in this wavelet expansion are called the discrete wavelet transform of the signal \( f(t) \). When the wavelet is orthogonal, these coefficient can be calculated by,

\[ c_j^k = \int_{-\infty}^{\infty} f(t) \varphi(2^j t - k) dt \]  

(2)

\[ d_j^k = \int_{-\infty}^{\infty} f(t) \psi(2^j t - k) dt \]  

(3)

Where \( \varphi(t) \) & \( \psi(t) \) are respectively, the scaling (approximation) & wavelet (detail) coefficient in the DWT. The frequency axis is divided into dyadic intervals towards the lower frequencies while the bandwidth length decreases exponentially[4].The set of wavelet defined a special filter bank which can be used for signal component analysis and resulting wavelet transform coefficient can be further applied signal features for its classification. The basic decomposition of the signal present in the Fig.2

![Decomposition of the Signal](image)

Fig. 3 Decomposition of the Signal.

**Characteristics of Wavelet**

Base wavelet are characterized by a number of properties that determine their applications in wavelet-based signal processing. Typical wavelet properties such as Orthogonality, symmetry & Compact Support.

1. **Orthogonality:** It means that the inner product of the base wavelet base wavelet with itself is unity & the innerproduct between the base wavelet & the scaled and shifted wavelet are zero.
   
   **Advantage:** The advantage of orthogonal wavelet is that the wavelet transform can enable signal decomposition into non-overlapping sub-frequency bands. High efficiency. Orthogonal wavelet are used for implementing the discrete wavelet transform and wavelet packet transform.

2. **Symmetry:** The symmetric property ensures that a base wavelet can serve a liner phase filter. This Is an important aspect for wavelet-based filtering operation as the absence of this property can lead to phase distortion.

3. **Compact Support:** It means that basic function of that wavelet is non-zero only on a finite intervals. This allows the wavelet transform to efficiently represent signals which have localized features. The efficiency of such representation is important for wavelet-based application such as data compression & signal detection.
4. Smoothness: This property determined by the no of vanishing moments. Recall that vanishing moments determines the smoothness of reconstruction. The dual vanishing moments determines the converge rate of multi-solution projection & necessary for detection signal antics.

Feature Classification

BCI Classifier’s Problem:

Some problem related to BCI classifier were reported by [7] which can be summarized as follows:

1. Course of Dimensionality and small training set: Dimensionality means that features come from various channels in different time intervals. So that when size of training dataset is small as compare to size of feature vector then classifier gives poor results.

2. Bias-Variance trade-off: It is due to the noise of BCI system i.e. biasness in mapping from & variation of training set that cause error. As a solution biasness & variation must be removed. By using classifier but the problem is that if we use stable classifier it has high biasness & low variance and if we use unstable classifier it has low biasness & high variance.

Support Vector Machine

An SVM also uses a discriminate hyperplane to identify classes. However, concerning SVM, the selected hyperplane is the one that maximizes the margins, i.e. the distance from the nearest training points. Maximizing the margins is known to increase the generalization capabilities. An SVM uses a regularization parameter C that enables accommodation to outliers and allows errors on the training set.

Such an SVM enables classification using linear decision boundaries, and is known as linear SVM. This classifier has been applied, always with success, to a relatively large number of synchronous BCI problems. However, it is possible to create nonlinear decision boundaries, with only a low increase of the classifier's complexity, by using the “kernel trick”. It consists in implicitly mapping the data to another space, generally of much higher dimensionality, using a kernel function \( K(x; y) \). The kernel generally used in BCI research is the Gaussian or Radial Basis Function (RBF) kernel:

\[
R(x, y) = \exp\left(\frac{-\|x - y\|^2}{2\sigma^2}\right)
\]

The corresponding SVM is known as Gaussian SVM or RBF SVM.

Multilayer Perceptron (MLP)

A Multilayer Perceptron is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate output. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training the network. MLP is a modification of the standard linear perceptron, which can distinguish data that is not linearly separable.
The basic neural network consists of 3 layers.

1. **Input layer**: The input layer consists of source nodes. This layer captures the features pattern for classification. The number of nodes in this layer depends upon the dimension of feature vector used at the input.

2. **Hidden layer**: This layer lies between the input and output layer. The number of hidden layers can be one or more. Each hidden layers have a specific number of nodes (neurons) called as hidden nodes or hidden neurons. The hidden nodes can be varying to get the desired performance. These hidden neurons play a significant role in performing higher order computations. The output of this layer is supplied to the next layer.

3. **Output layer**: The output layer is the end layer of neural network. It results the output after features is passed through neural network. The set of outputs in output layer decides the overall response of the neural network for a supplied input features.

**Strengths and weaknesses of Classifiers:**[7][8]

1. **SVM**:
   - **Strengths**:
     - Provides maximized margin (degree of separate) in training data.
     - Due to use of kernel SVM is flexible in threshold attributes.
     - Best in classification and always provides single unique solution rather than multiple solution.
     - Generally best overall accuracy as compare to KNN and ANN.
     - It did not require training again and again.
     - Best for those BCI features which contains noise & oatiler.
     - Low error rate.
     - BCI EEG data contain high dimensionality therefore SVM gives good results in case of high dimensionality and smaller training set
   - **Weakness**:
     - In testing phase SVM is slow.
     - Contain high complexity = O(n²).
     - Wastage of memory.
     - If dataset is large then no solution to train it.

2. **MLP**:
   - **Strengths**:
     - MLP d not use assumption or guess for decision making, it is capable to provide decision directly from trained data.
     - Capable to learn complex decision boundaries.
     - Good in classification performance but not good for discriminating the non class data.
     - Best in signal classification.
     - Capable of reduce signal noise.
   - **Weakness**:
     - High complexity i.e. O(n²).
     - Only provides solution against linearly separated data.
     - Sensitive to overtraining especially with such noisy and non-stationary EEG data.
     - Recognition.
### Performance analysis of classifiers:

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Computation Complexity</th>
<th>Accuracy</th>
<th>Speed</th>
<th>Memory Need</th>
<th>BCI Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>Initially high i.e. $O(n^2)$ but becomes less due to kernel tricks</td>
<td>Best</td>
<td>Classification = Fast, Testing = slow, Exhaustive hypertraining = slow, Overall luckily very fast in BCI application.</td>
<td>Large</td>
<td>Yes</td>
</tr>
<tr>
<td>MLP</td>
<td>$O(n^2)$</td>
<td>Low learning accuracy</td>
<td>Good in classification, slow in discrimination</td>
<td>??</td>
<td>Yes applicable to all BCI</td>
</tr>
<tr>
<td>LDA</td>
<td>High computation and structure complexity</td>
<td>- In low dimensional feature = High, - In high dimensional feature = slow</td>
<td>Classification = Slow, Training = fast</td>
<td>??</td>
<td>Yes for Motor imaginary BCI</td>
</tr>
<tr>
<td>HMM</td>
<td>High time complexity</td>
<td>Very accurate in time series &amp; raw EEG data.</td>
<td>Fast in execution time</td>
<td>Large</td>
<td>Yes applicable for all BCI</td>
</tr>
</tbody>
</table>

**4. FUTURE WORK**

In the proposed project we tested the result for detection of the Parkinson’s disease by using hybrid classifier which is the combination of support vector machine classifier and multilayer perceptron classifier. Then we can compare the results obtained by this work with the system which is formed by using single classifier (i.e. system implemented using SVM, and the system implemented using MLP).

In this work 8 channel signals are used for the human computer interface. The process would become easier and more accurate if 16 channels are used for analysis and implementation.
5. CONCLUSION

Brain computer Interface is the simplest method to interface with any real input and output device using Electroencephalograph (EEG) signals based on the frequency, the signal can be classified into different bands of frequency. Each band will predict different conditions with the help of these frequencies control of interfaced device can be made easy and automated. In the proposed method the wavelet transform is used for the feature extraction, because of the advantages of this on time domain analysis and frequency domain analysis. By using wavelet we can get the signal representation in time-frequency domain. Instead of using one classification algorithm, we are combining the two classifier such as Support vector machine (SVM) & Multilayer Perceptron (MLP) because the SVM has best training accuracy and the MLP has best testing accuracy than other.

6. REFERENCES

[1]. P.V.RamaRaju, V.M.Rao, “Relevance of wavelet transform for EEG signals”.
[3]. Dr.D.S.Bormane and Prof.S.T.Patil, “Pearl ensemble classifier Decision”.
[5]. Aleas Prochazka, Jaromair Kul Oldrich Vysata, “Wavelet Transform use for feature extraction & EEG signal classification.
[6]. Chia-Jung Chang, “Time frequency analysis and wavelet transform tutorial”, Time frequency analysis for biomedical engineering, National Taiwan University.
[10]. Zahari, Nooritawati, Ihsan, “Classification of Parkinson’s Disease based on Multilayer Perceptron neural Network”