CLASSIFICATION OF ELECTROENCEPHALOGRAPHY SIGNALS USING MIXTURE OF FEATURES

A THESIS SUBMITTED IN PARTIAL REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF TECHNOLOGY
IN
ELECTRONICS & COMMUNICATION ENGINEERING
BY

PANKAJ KUMAR SANGRA
ROLL NO. 107EC023

UNDER THE GUIDANCE OF

Dr. SAMIT ARI
ASSISTANT PROFESSOR
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA
ORISSA 769 008
INDIA

National Institute of Technology
Rourkela

CERTIFICATE

This is to certify that the thesis entitled, "Classification of Electroencephalography signals using mixture of Features" submitted by Pankaj Kumar Sangra in partial fulfillment of requirements for the award of Bachelors of Technology degree in Electronics and Communication Engineering, Department of Electronics and Communication Engineering at National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any University/Institute for award of any degree or diploma.

Date: 13th May, 2011

Dr. Samit Ari
Assistant Professor
Dept. of Electronics & Communication Engg.
National Institute of Technology Rourkela
Orissa- 769008

ACKNOWLEDGEMENT

I have this opportunity to thank all the concern individuals whose guidance, help and timely support made me to complete this project within the stipulated time. Unfortunately, it is not possible to express my thanks to all of them in a single page of acknowledgement.

First and foremost I would like to express my sincere gratitude to my project supervisor Dr. Samit Ari for his great support, nice motivation, proper guidance and constant encouragement which made this project enrich and insightful experience. His timely inputs, valuable return and constructive criticism in many stages of this work was act as catalyst in understanding and execution of this project.

I am also grateful to Prof. S.K. Patra (Head of the Department), Department of Electronics and Communication Engineering for assigning me the real life application based project and providing me with various facilities of the department. His valuable suggestions and inspiring guidance was extremely helpful.

An assemblage of this nature could never have been attempted without reference to and inspiration from the works of others whose details are mentioned in reference section. I acknowledge our indebtedness to all of them.

I would also like to thank all professors and lecturers, and members of the department for their generous help in various ways for the completion of this thesis. I also extend our thanks to our dear friends for their support and cooperation.

Date: 13th May 2011
Place: NIT Rourkela

Pankaj Kumar Sangra
Dept. of ECE
National Institute of Technology Rourkela
Orissa-769008

Contents

ABSTRACT

Electroencephalography (EEG) signals provide valuable information to study the brain function and neurobiological disorders. Digital signal processing gives the important tools for the analysis of EEG signals. The primarily focus on classification of EEG signals using different feature extraction methods for pattern recognition purpose. The various tools are used for extracting the relevant information from EEG data is Discrete Wavelet Transform (DWT), Spectral analysis using Autoregressive (AR) Model and Lyapunov Exponents. The EEG data was collected from standard repository source. The two classifiers ANN and CNN are used for the classification purpose. A technique is proposed based on using the combined features extracted from different methods. In committed neural network, several independent neural networks are trained by the extracted features from different EEG signals are constituted a committee. This committee takes the final decisions for classification which in turn represents a combined response of the individual networks. The performance of the proposed algorithm is evaluated on 300 different recordings from three different cases comprising of healthy volunteers with eyes open, epilepsy patients in the epileptogenic zone during a seizure-free interval, and epilepsy patients during epileptic seizures. The experimental results show that the classification performance for the proposed technique is higher than some of the earlier established techniques.

Keywords: EEG signal, Discrete wavelet transform (DWT), Autoregressive (AR) Model, Lyapunov expoenents, Artificial Neural Network (ANN), Committee neural network (CNN)

List of Figures

List of Tables

List of Abbreviations

EEG	Electroencephalography
CNS	Central Nervous System
EMG	Electromyogram
ADC	Analog to Digital Converter
WT	Wavelet Transform
STFT	Short Time Fourier Transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
DWC	Discrete Wavelet Coefficients
FFT	Fast Fourier transform
AR	Autoregressive

PSD	Power Spectral Density
MA	Moving Average
ARMA	Autoregressive Moving Average
ARC	Discrete Wavelet Coefficients
LE	Lyapunov Exponent
ANN	Artificial Neural Network
MLPNN	Multilayer Perceptron Neural Network
NN	Neural Network
NN1	First Neural Network
NN2	Second Neural Network
NN3	Third Neural Network
CNN	Committee Neural Network

CHAPTER 1
INTRODUCTION

Electroencephalography (EEG) is the recording of brain activity along the scalp in the form of electrical signals. The term EEG refers that the brain activity emits the signal from head and being drawn. It is produced by bombardment of neurons within the brain [1]. It is measured for a short duration of 20-40 minutes with the help of placing multiple electrodes over the scalp [2]. The researchers have found that these signals indicate the brain function and status of whole body. So, the EEG signals are mainly used in clinical laboratories for the diagnosis of epilepsy diseases as the actual EEG signals get distorted [3]. This ability motivates us to apply the advanced digital signal processing techniques on the EEG signals. The study of neuronal function and neurophysiological properties with EEG signals generation and their detection plays the vital role in detection, diagnosis and healing of brain disorders and brain diseases.

1.1 Past History of EEG signal

In 1818-1896, the first electrical signals noticed from muscle nerves were measured using galvanometer and termed the concept called neurophysiology [4], [5], [6]. Initially the two electrodes were placed over the scalp and the first brain activity measured in the form of electrical signals in 1875 [4].

It was first measured from dog’s brain then later it started to measure EEG signals from human brain. The first recording was made of 1-3 minutes on a photographic paper. The discoverer Hans Berger concluded by his research that the major part of EEG signals consists of alpha rhythms [4], [7].

1.2 Neural Activities

The *Central Nervous System (CNS)* consists of two types of cells, nerve cells and glia cells. All the nerve cell consists of axons, dendrites and cell bodies. The proteins developed in the cell body are transmitting information to other parts of the body. The long cylindrical shaped axon transmits the electrical impulse. Dendrites are connected either to the axons or dendrites of other inside cells and receive the electrical impulse from other nerves cells. The study of brain found that each nerve of human is approximately connected to 10000 other nerves [4].

The electrical activity is mainly due the current flow between the tip of dendrites and axons, dendrites and dendrites of cells. The level of these signals is in mV range and its frequency is less than 100Hz [4].

1.3 EEG Generation

The EEG signal is the current measured between the dendrites of nerve cells in the cerebrum region of the brain during their synaptic excitation. This current consists of electric field detected by electroencephalography (EEG) equipment and the magnetic field quantified by electromyogram (EMG) devices [4].

The human head consists of many parts some of which is scalp, skull, and brain. The noise is either to be produced internally in the brain or externally over the scalp due to the system used for recording. The degree of attenuation caused by skull is nearly hundred times larger than the soft tissues present in the head. As the level of signal amplitude is very low, the more number of neuron in excitation state can only produce the recordable potential by the scalp electrode system. Thus the amplifiers are used to increase the level of signals for later processing [8].

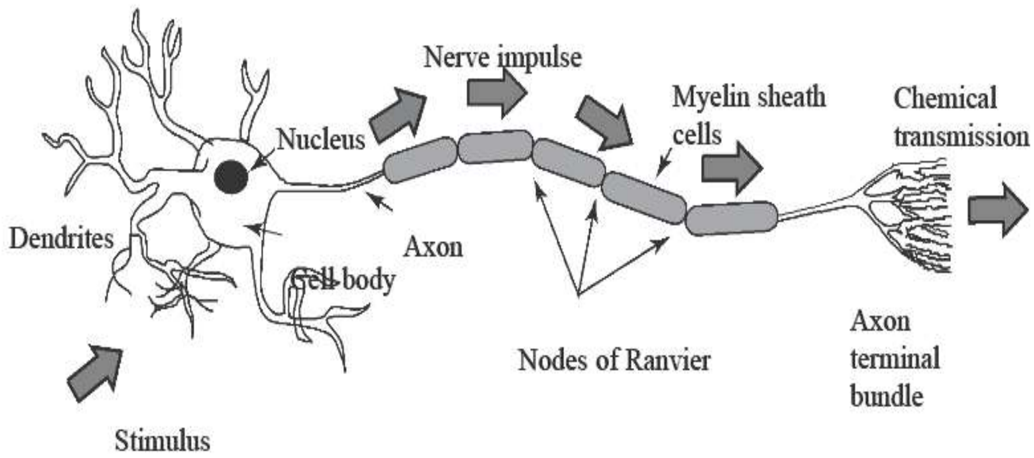


Fig. 1.1 Structure of neuron [4], [8]

The brain structure is divided into three regions, cerebrum, cerebellum and brain stem. The cerebrum region defines the initiation of movement, conscious sensation and state of mind. The cerebellum region plays a role in voluntary actions like muscle movements. The brain stem region controls the respiration functioning, heart regulation, biorhythms and neural hormones [4], [9]. It is clear that the EEG signals generated from brain can determine the status of whole body and brain disorders. [4], [9], [10].

1.4 EEG Recording

The first electrical variation was noted by using a simple galvanometer. But in recent generation they are recorded by using EEG systems which consist of multiple electrodes, amplifiers single for each channel to amplify the attenuated signals and followed by filters to remove the system noise and registers [4], [9]. Later, EEG systems are equipped with computerized system to store the large data, for easy and correct analysis of EEG signals by using high sampling rate and more number of quantization levels with the help of advanced signal processing tools.

The Analog to Digital Converter (ADC) circuits are used to convert the analog EEG signal to digital form. The bandwidth for EEG signal is 100 Hz. So, the minimum 200 samples/sec are required for sampling the EEG signal to satisfy the Nyquist criterion. Sometimes, the higher sampling rate of 2000 samples/sec is used for getting high resolution of EEG signals. The capacity of connected memory units depends upon the amount of data recorded. The different types of electrodes used for recording high quality data.

1.4.1 Conventional Electrode Positioning

The International Federation of societies for Electroencephalography and Clinical Neurophysiology has recommended the standard electrode system for 21 electrodes which is also called as 10-20 electrode setting [11].

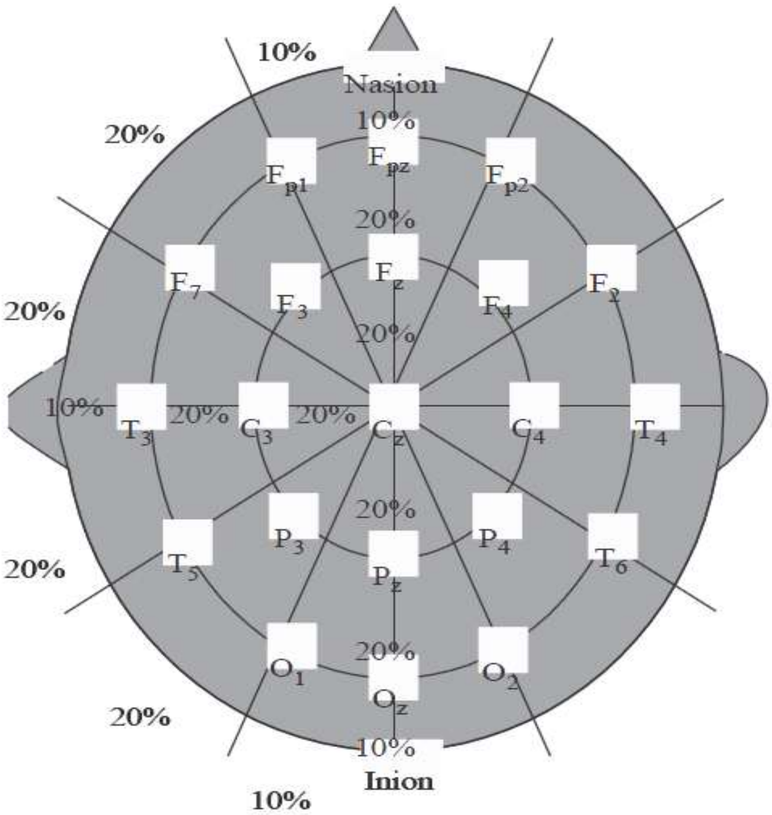


Fig. 1.2 Conventional 10 – 20 EEG electrode positions for the placement of 21 electrodes [4]

1.5 Brain Rhythms

The shape of EEG signal also identifies the most of the brain disorders. The shape of EEG signals depends upon human age, health and its state such as sleep and awake. The brain waves is divided into five types each having different frequency ranges are alpha (α), beta (β), delta (δ), gamma (γ) and theta (θ).

1. *Alpha*: These waves are found in the posterior side and appear over the occipital region of the brain. The frequency of these waves lies in between 8-13 Hz. These are mostly in round or sinusoidal shaped. These waves are introduced when the person is in pure relaxed state or with closed eyes. The origin of alpha waves is not clear and this topic is still under the research [4], [12].
2. *Beta*: The frequency variation in these types of signals lies in between 14-26 Hz. It is usually produced by adults in waking state when their brain is under highly attentive, thinking state, focusing and concentrating on heavily tasks. A high amplitude signal may arise when the person is in panic situation. These are mostly appeared in frontal and central portion of the brain. The amplitude of beta wave is less than 30 μV.
3. *Gamma*: The frequency range of these waves falls from 30-45 Hz. It is also called as fast beta waves. These waves have very low amplitude and present rarely. But, the detection of these rhythms plays an important role in finding the neurological diseases. These waves are occurred in front central part of the brain. It suggests the event-related synchronization (ERS) of the brain [4], [13].
4. *Delta*: These waves have the frequency range 0.5-4 Hz which is lower than the alpha band. The deep sleep is the primarily source of these kind of waves. It starts from the deep inside the brain and attenuates by a large amount due to skull [4].
5. *Theta*: These waves are assumed to be originated from thalamic region within the brain. The frequency components of delta waves lie in between 4-4.75 Hz. The

frequency of alpha waves decreases gradually with the duration of time in healers and expert mediators. The feelings and maturational are observed with the changes in rhythm of waves [4], [14].

The frequency band higher than the EEG signal lies in the range of 200-300 Hz which is found in cerebellar region the animals. [15], [16].

1.6 Applications

Clinical use

The EEG signal is used in the clinics for the diagnosis of following problems [2], [9], [10].

i) EEG signals are used for regulating the depth for anesthesia.

ii) EEG signals are used for determining epilepsy and source point of seizure.

iii) EEG signals are used for monitoring the brain attentiveness, coma brain death and brain development.

iv) EEG signals are used for identification of sleep disorders, brain disorders and physiology.

The use of EEG signals in medical field applications attract the researchers towards EEG signal analysis as their research topic. The EEG signal analysis can be performed using advanced digital signal processing tools.

Research use

EEG is widely used in various fields [2] such as

i) Cognitive science

ii) Cognitive psychology

iii) Neuroscience

iv) Psychophysiological

1.7 Data Base

The raw EEG signal is obtained from repository source as mentioned in [17] which consists of total 5 sets (classes) of data (SET A, SET B, SET C, SET D and SET E) corresponding to five different pathological and normal cases. Three data sets are selected from 5 available classes of data for the present study. The three types of data represent three class of EEG signals (SET A contains recordings from healthy volunteers with open eyes, SET D contains recording of epilepsy patients in the epileptogenic zone during the seizure-free interval, and SET E contains the recordings of epilepsy patients during epileptic seizures). All recordings were measured using Standard Electrode placement scheme also called as International 10-20 system. Each data set contains the 100 single channel recordings. The length of each single channel recording was of 26.3 sec. The 128 channel amplifier had been used for each channel [18]. The data were sampled at a rate of 173.61 samples per second using the 12 bit ADC. So the total samples present in single channel recording are nearly equal to 4097 samples (173.61*23.6). The band pass filter was fixed at 0.53-40 Hz (12dB/octave) [19] [20].

1.8 Raw EEG Signal

The waveform of single channel recording of three different types of EEG signals (SET A, SET D, and SET E) is shown in Fig. 1.3. As it is already explained that each EEG signal contain 4097 samples. The picture itself dictates that three signals are different.

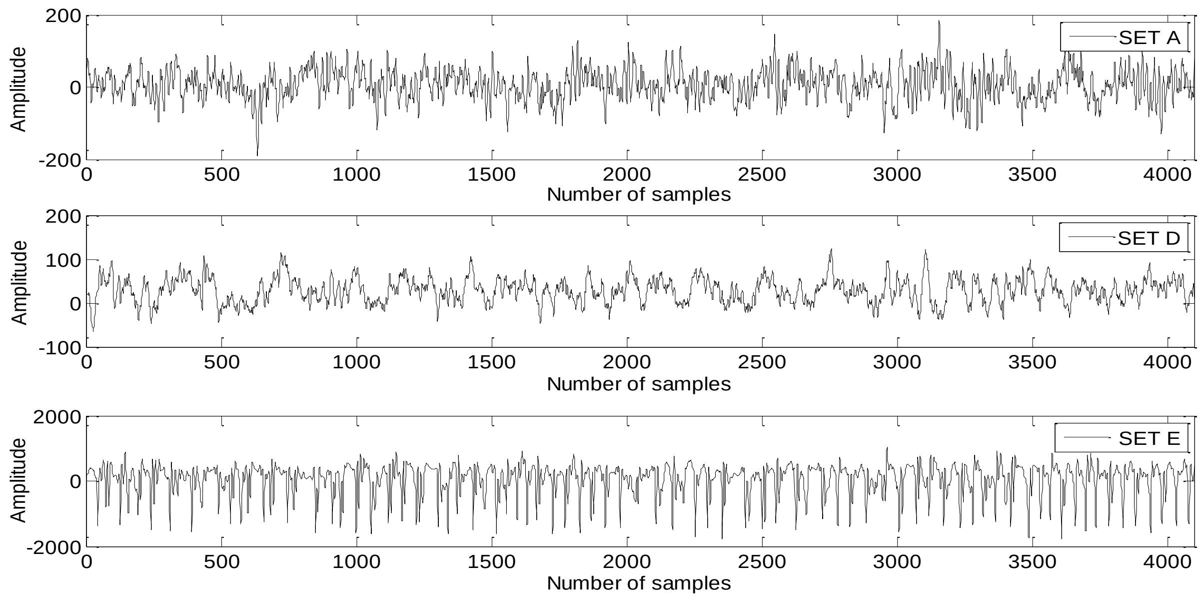


Fig. 1.3 Waveform of raw EEG signals of Cass A, Class D and Class E

1.9 Objective

The main objective of our research is to analyze the acquired EEG signals using signal processing tools and classify them into different classes. The secondary goal is to improve the accuracy of classification. In the present research, the different feature extraction methods have been used for pattern recognition problem. The mixture of features used for classification purposes is purposed in this research. Only the important features are input to the neural networks for training and testing purposes.

The four successive stages followed in the characterization of EEG signals are: normalization, segment detection, feature extraction and classification. All these processes are shown in the flow diagram of Fig. 1.4.

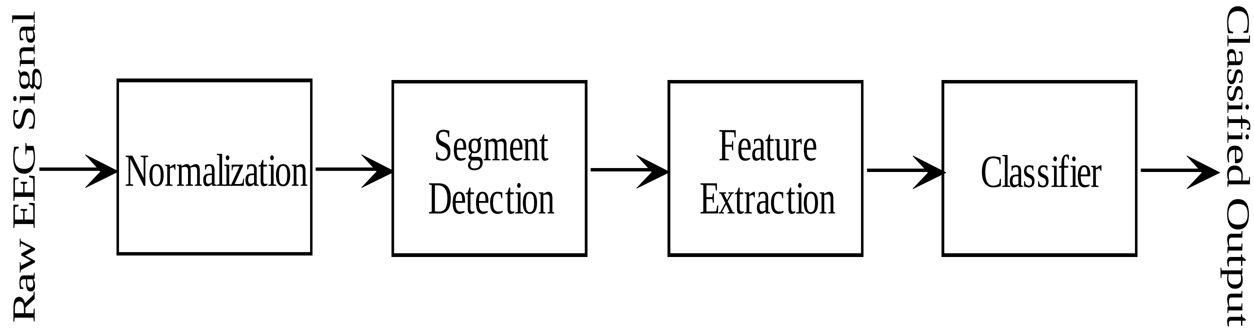


Fig. 1.4 Systematic block diagram for classification of EEG signals

CHAPTER 2
FEATURE EXTRACTION TECHNIQUES

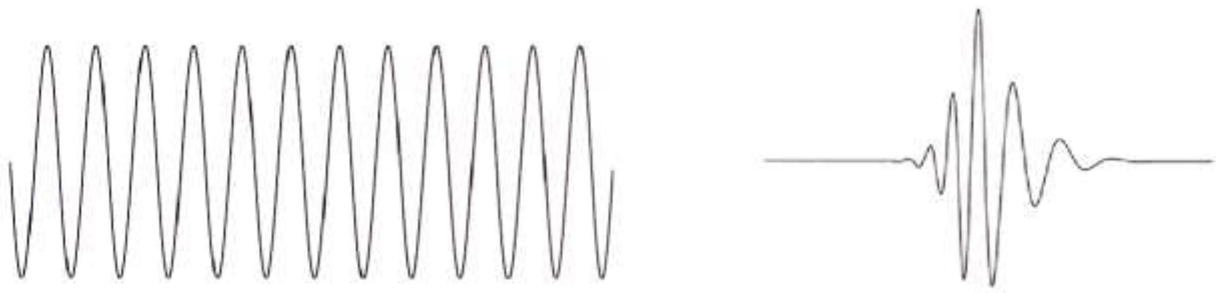
Feature extraction is the collection of relevant information from the signal. Due to high transitions the signals are unable to distinct and insufficient to answer the status of individual. So, the DSP tools have been used for extracting the desired characteristics of different EEG signals and provided the knowledge of human state. A number of feature extraction methods have been present for study the EEG signals.

2.1 Wavelet Transform

The signal can also be represented to another form by its transform by keeping the signal information to be same. The *Short Time Fourier Transform (STFT)* is used to study the non-stationary signals. To overcome the limit of its use to only non-stationary signals, the wavelet transform was developed to represent the signal into a function of time and frequency [21]. The STFT provides the identical resolution at all frequencies, but the use of wavelet transform provided a significance advantage to analyze the different frequencies of the signals with different resolutions by the method called multi-resolution technique [21]. So, this is perfect for analyzing the EEG signals [22].

The wavelets are the localized waves as their energy is concentrated in time which makes it best suitable for the analysis of transient signals [21].

The process of wavelet analysis is similar to that of STFT. In the first step of wavelet analysis, original signal is multiplied with the wavelet same as the signal is multiplied with window function in STFT case, and in the second step the transform is computed for each individual segment.



(a) (b)

Fig. 2.1 Shape of (a) wave, (b) wavelet [21]

The size of wavelet varies depending upon the frequency components in wavelet transform which provides time or frequency resolution at all frequencies. There is a tradeoff between time resolution and frequency resolution. The Wavelet transform provides the good time resolution with poor frequency resolution at high frequency and good frequency resolution with poor time resolution at low frequency [21].

2.1.1 Continuous Wavelet Transform

The wavelet transform only finds the convolution between signal and basis function. The wavelet or basis functions used in the wavelet transform are derived from mother wavelet by scaling and translation process [21]. The equation of *Continuous Wavelet Transform (CWT)* is defined as

$$X_{CWT}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t).\Psi(\frac{t-\tau}{s})dt \tag{2.1}$$

where, $x(t)$ is the raw signal and $\Psi(t)$ is the mother wavelet.

The basis functions are to be generated based on desired characteristics. The τ is the translation parameter which defines the location of wavelet function for time information in wavelet transform. The scale parameter s corresponds to frequency information and it is defined as |1/frequency|. The large value of scale parameter (means large time scale, low frequencies) expands the signal and provides the deep detail of information hidden in the signal whereas the small scale parameter (less time duration, high frequencies) compresses the signal and provides the overall information about the signal [21]. High frequencies don't stay for long time but low frequencies stay for long duration of time. Hence, this approach of analysis is mostly used in practical applications.

2.1.2 Discrete Wavelet Transform

The wavelet series is obtained by sampling the signal and represent the discretized CWT [21]. It helps in easier computation of CWT by using computers. The sampling frequency can varies with the change of scaling without violating the Nyquist rule. The *Discrete Wavelet Transform (DWT)* results the faster computation of wavelet transform by using sub-band coding method. It is easier in implementation, less computational time and uses fewer resources.

In CWT, the signal is analyzed by using basis function but in DWT, the digital signal is represented in time scale by using digital filtering technique [21]. The signal with different frequency components is passed through the filters of different cutoff frequencies at different scales.

The operation of DWT is performed by using multiple filters with rescaling. The filters are used to measure the detail information about the signal and its scale is determined by decimation and interpolation process [23], [24].

The successive decomposition of the discrete signal into low and high frequencies is shown in Fig. 2.2 where $x[n]$ represents the discrete signal and $n \varepsilon I$ denotes the number of sample. The low and high pass filters are denoted by $G[n]$ and $H[n]$ respectively. The high pass filters allows only the high frequency components [21] and produces the detail coefficients $d[n]$. The low pass filter allows only low frequency components to pass and produces the approximate coefficients $a[n]$. These approximate coefficients are further allowed to decompose in a successive manner and produce detail coefficients at each level of decomposition [23], [24]. The Fig 2.2 is shown with the 4 level decompositions.

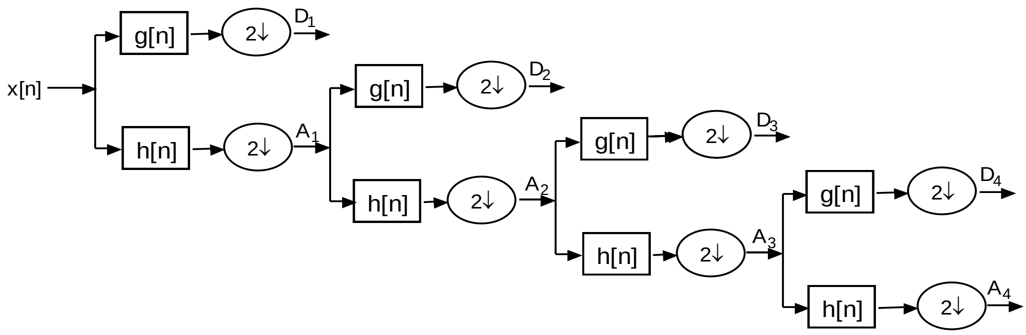


Fig. 2.2 Four level Wavelet decomposition [25]

At each level of decomposition, In half band high pass filtering, the frequency resolution doubles by a decimation factor 2 and halves the times resolution as it produces the signal with half the samples without any loss of data while in the half band low pass filtering, the frequency resolution reduced by 2 as half of the frequencies are removed by which it doubles the time resolution by decimation of factor 2. Thus the multi resolution approach produces good time resolution at high frequencies and good frequency resolution at low frequencies [21]. The DWT of the signal is obtained by concatenating the detail coefficients and approximate coefficients starting from the end, i.e. $a4d4d3d2d1$.

2.1.3 Wavelet Families

The basis function is used as a mother wavelet for obtaining wavelet transform and it can be produced through translation and scaling the mother wavelet. As the mother wavelet produces all the basis functions which determine the characteristics of resultant wavelet transform, the selection of appropriate mother wavelet is necessary for finding of the wavelet transform effectively. The some of the commonly used wavelet functions are shown in Fig. 2.3.

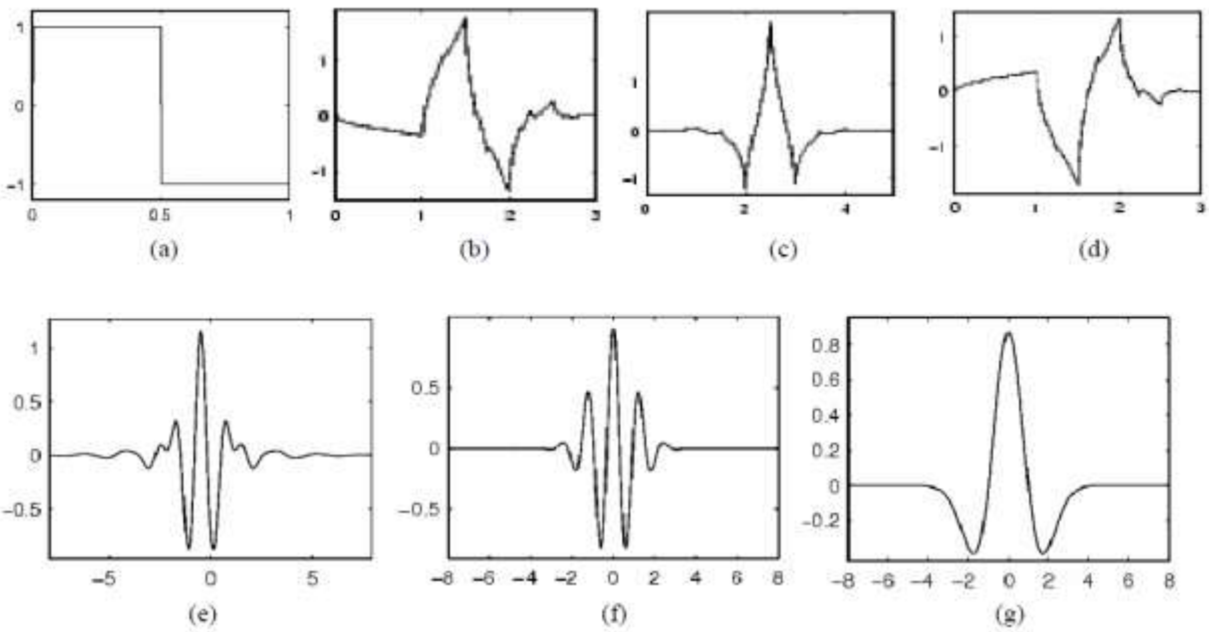


Fig. 2.3 Wavelets functions (a) Haar (b) Daubechies (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat [21]

The Daubechies wavelet is important one and is used in various applications. The capability of wavelet to analyze the signal is different and it can be chosen based upon its shapes

2.2 Autoregressive Method

The power spectral density shows the distribution of power with frequency. The spectral analysis of EEG signal can be performed by three methods- *Burg autoregressive*

(AR) moving average (MA) and modified Yule Walker *autoregressive moving average (ARMA)* [26][27], [28], [29]. These method are based on modeling the output data sequence *x(n)* described by a rational system. These models based methods estimates the spectrum in two steps- first step involved the estimation of process parameters but in the second step, this estimation is used for computing the PSD [29]. This method is widely used due to the reason that the parameters are estimated easily by just simplifying the linear equations. These models also provide good stability for small length signals and good spectral resolution [29]. Power spectral density, R_{xx} of the random stationary signal can be rewritten in terms of the polynomials $A(z)$ and $B(z)$ whose roots fall inside the unit circle in the z-plane as shown by the formula [23].

$$R_{xx}(z) = \sigma_w^2 \frac{B(z)B(z^{-1})}{A(z)A(z^{-1})}, \qquad r_1 < |z| < r_2 \tag{2.2}$$

where, σ_w is the variance of the white Gaussian noise $w(n)$.

Then the linear filter $H(z)$ for generating the random process $x(n)$ from the white noise samples $w(n)$ is written as

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^{q} b_k z^{-k}}{1 + \sum_{k=1}^{p} a_k z^{-k}}, \qquad |z| > r_1 \tag{2.3}$$

where, a_k and b_k are the filter coefficients which locates the position of the poles and zeros of $H(z)$, respectively.

For the linear system with the rational system function $H(z)$ given by above relation [26], the difference equation between the input $w(n)$ and output $x(n)$ is shown below

$$x(n)+\sum_{k=1}^{p}a_{k}x(n-k)=\sum_{k=0}^{q}b_{k}w(n-k) \tag{2.4}$$

In case of *Autoregressive (AR) process*, $b_0=1,\ b_k=0,\ k>0.$ Then the linear filter is an all pole filter [26], i.e. $H(z)=1/A(z)$ and the reduced difference equation for AR model is

$$x(n)+\sum_{k=1}^{p}a_{k}x(n-k)=w(n) \tag{2.5}$$

In case of *moving average (MA) process*, $a_k=0,\quad k\geq 1.$ Then the linear filter is an all zero filter i.e. $H(z)=\ B(z)$ and the difference equation for the moving average (MA) process can be written as

$$x(n)=\sum_{k=0}^{q}b_{k}w(n-k) \tag{2.6}$$

In case of *autoregressive, moving average (ARMA) process*, the linear filter will be $H(z)=B(z)/A(z)$ which is taken from [26]. It has both finite poles and zeros in the z-plane and the difference equation can be expressed as

$$x(n)+\sum_{k=1}^{p}a_{k}x(n-k)=\sum_{k=0}^{q}b_{k}w(n-k) \tag{2.7}$$

The power spectral density (PSD) of the EEG signals is computed by using Burg Autoregressive (AR) Model in the present work. This method is based on minimization of forward and backward prediction errors while constraining the AR parameters to satisfy the Levinson-Durbin recursion process [28], [30] The Burg's method is a recursive method.

Autoregressive coefficients provide us the important features in terms of the power spectral density (PSD). The Burg method estimates the reflection coefficients a_k. Since, this method describes the input signals by using the all pole model. So the selection of model order is critical because the very low model order produces smooth spectrum and too large model order effect in stability. Any stochastic process can be modeled by using AR process.

2.3 Lyapunov Exponents

This method is widely used in medical field for finding the relevant information from the signals [31], [32]. Therefore the *Lyapunov exponents (LE)* are used here for the extraction of features from EEG signals.

The Lyapunov exponents describe the qualitative behavior of the dynamical systems. The Lyapunov exponents can be easily calculated if we know the system behavior [33]. The number of Lyapunov exponents depends upon the number of state variables of the system [33]. For the one-dimensional systems only largest Lyapunov exponent can be computed, whereas in multidimensional systems we can compute the whole Lyapunov spectra.

There are two methods for estimation of Lyapunov exponents. One is called direct method where the Lyapunov exponents are calculated from a time series without knowing the system behavior. We calculate it by forming the reconstruction matrix [33].

Lyapunov exponents are used to find the stability of the steady state systems. The chaotic behavior of the system is due to the reason that the phase space trajectories which have close initially states and will separates gradually [34], [35], [36], [37], and [18].

The Lyapunov exponent defines the rate at which two trajectories will separate from each other exponentially [41]. It is defined by the relation

$$\frac{\| \delta x_i(t) \|}{\| \delta x_i(0) \|} = 2^{\lambda_i t} (t \to \infty)$$
$$\text{or} \tag{2.8}$$
$$\lambda_i = \lim_{t \to \infty} \frac{1}{t} \log_2 \frac{\| \delta x_i(t) \|}{\| \delta x_i(0) \|}$$

where $\| \delta x_i(0) \|$ and $\| \delta x_i(t) \||$ represents the distances of the points at the i^{th} direction which were close initially. The positive Lyapunov exponents show the chaos nature [34], [35], [36], [37], [42]. This shows that the initially closed points will separate abruptly with iterations in the i^{th} direction [18]. This phenomenon sometimes is known as sensitive dependency on initial states. There are lots of methods which have been already developed for finding the Lyapunov exponents [38].

Mostly the Lyapunov exponents can be estimated either from the differential equations [32] of the dynamic system [39] or from an experimentally observed time series [40]. There have been two approaches for computing the Lyapunov exponents from an experimental data. The first approach is based on the evolution of neighboring points in space [41], [18]. This method provides only the single and largest Lyapunov exponent. The second one is based on computation of local Jacobi's matrix [42], [18]. This method evaluates the full spectrum Lyapunov exponents [43], [44], [32]. The local Jacobi's matrix algorithm is used in our work for computing the Lyapunov exponents. The Lyapunov exponent method is easy to implement, faster and robust to the changes in the embedding dimension, size of data sets, reconstruction delay and noise level.

2.3.1 Algorithm for finding the Lyapunov exponents:

We have the observed time series of EEG signal. This procedure shows how we can find the Lyapunov exponents from a given time series.

Let $\{x_1, x_2, x_3, \ldots, x_N\}$ be the time series of length *N*.

Step 1: Reconstruct the attractor dynamics from a time series [45], [42].

The reconstructed trajectory *X* can be represented in a matrix; the each row of it denotes the phase space vector, i.e.

$$X=[X_1\ X_2\ X_{3\ldots}X_M]^T$$

where, X_i denotes the state of the system at discrete time *i*. For *N* point observed series, each X_i is written as

$$X_i = [x_i, x_{i+\Delta t}, x_{i+2\Delta t}, ..., x_{i+(m-1)\Delta t}]$$

Where,

Δt = lag

d = the embedding dimension

The order of matrix $X = M*d$.

The embedding dimension *d* is calculated from the Taken's theorem ($d>2n$; *n* is the number of Lyapunov exponents). The lag Δt can be calculated by using fast Fourier transform (FFT).

The relation between the constants *m*, *M*, Δt and *N* is given by [45]

$$M=N-(d-1)*\Delta t$$

The reconstructed trajectory will be formed where each row is the orbital point in phase space.

Step 2: Start searching the orbital points included in the ball of radius ϵ by computing the distance y^i [42]. It is represented as

$$\{y^i\} = \{X_{k_i} - X_j \mid \left\| X_{k_i} - X_j \right\| \le \varepsilon\},$$

Searching all the points is the time consuming process so we choose the upper cutoff limit.

First we take the orbital point as the reference point and start searching the points which are included in the radius ϵ by considering reference point at the center of the ball. While searching if the number of points included the ball exceeds the upper limit then we stop the search and follow till step 4. If it is not exceeds, then check the condition if number of inside points is greater than equal to *d* (means *n*>=*d*) and all the search points (X_k) get exhausted, then follow the same steps till step 4 otherwise skip that reference point and go to the next point. Consider the next point as reference point and repeat the same process up to the end orbital point.

Let $\{X_{k_i}\}(i = 1, 2, 3, 4 \dots N)$ be the number of points lies inside the ball then finding the displacement vector [42] y^i between X_{k_i} and X_j is

$$\{y^i\} = \{X_{k_i} - X_j \mid \left\| X_{k_i} - X_j \right\| \le \varepsilon\},$$

where $\|W\|$ is the Euclidean norm which is defined by

$$\|W\| = \left(w_1^2 + w_2^2 + w_3^2 + \cdots + w_d^2 \right)^{\frac{1}{2}}.$$

Step 3: After the evolution of time $\tau = m\Delta t$ .The X_j will move to X_{j+m} and neighboring points $\{X_{k_i}\}$ will move to $\{X_{k_i+m}\}$ [42]. Then again start searching the point included in the ball by taking X_{j+m} as the reference point. Then we calculate the $\{z^i\}$ by the equation given below

$$\{z^i\} = \{X_{k_i+m} - \mathrm{X}_{j+m} \mid \left\| X_{k_i} - X_j \right\| \le \varepsilon\},$$

The parameter $\{y^i\}$ is mapped to $\{z^i\}$ and it can be written as $z^i = A_j y^i$

Step 4: This step finds the optimal estimation of A_j .We can evaluate it by solving the matrix equation given as

$$A_j V = C$$

Where, $$(V)_{kl} = \frac{1}{N}\sum_{i=1}^{N} y^{ik} y^{il} \quad \text{and} \quad (C)_{kl} = \frac{1}{N}\sum_{i=1}^{N} z^{ik} y^{il}$$

where the symbols y^{ik} and z^{ik} represents the k^{th} element of y^i and z^{ik} .

Steps 5: The basis vectors to be formed from the *Aj* matrix [42] and λ_i which explains the rate of

convergence or divergence can be calculated from the formula :

$$\lambda_i = \lim_{n \to \infty} \frac{1}{n\tau} \sum_{j=1}^{n} \left\| A_j e_i^j \right\|$$

Where $i = 1, 2, 3, 4 \ldots d$ and e_i^j is the basis vectors of A_j. The *d* number of Lyapunov exponents is obtained for *d* iterations.

CHAPTER 3
CLASSIFICATION OF EEG SIGNALS

The classification of EEG signals is highly important in medical field. This can provide the knowledge of status of human mind. This helps in curing many diseases after diagnosis or classification results. So this is the important step carried out after the feature extraction for analyzing the EEG signals. There are many signal processing algorithms present for studying the EEG signals. We have used only two algorithms for classification and also known as *classifiers*.

3.1 Artificial Neural Network (ANN)

Artificial neural network (ANN) has been the successfully used classifier in numerous fields. So, it is of interest to use it for EEG analysis. It can be modeled on a human brain [30]. The basic processing unit of brain is neuron which works identically in ANN. The neural network is formed by a set of neurons interconnected with each other through the synaptic weights. It is used to acquire knowledge in the learning phase. The number of neurons and synaptic weights can be changed according to desired design perspective [37], [30]. 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 [37], [30].

The typical structure of neural network is shown in Fig. 3.1 which consists of m input neurons in general and n hidden neurons with single hidden layer. The output layer has only three neurons. The network is called as fully connected network when all the neurons are connect with the adjacent neurons [30].

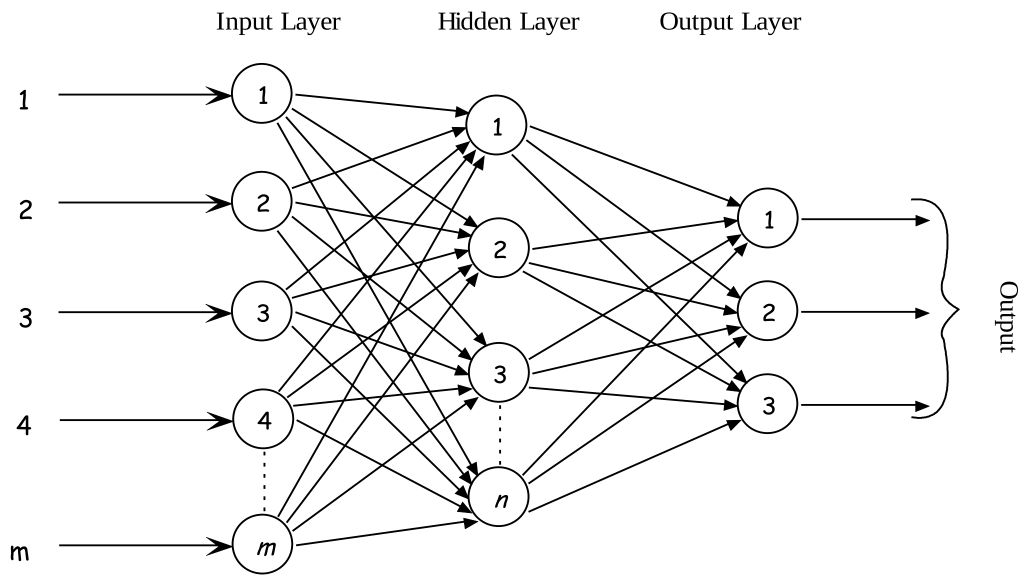


Fig. 3.1 Structure of neural network

3.2 Multilayer Perceptron Neural Network (MLPNN)

The ability of classifying the non-linearly separable classes in a supervised manner enable it mostly used neural network for classification purpose. It uses the error correction rule known as *back propagation algorithm* [37], [30].

This principle of this algorithm is based on delta rule and gradient descent [30]. In the delta rule, the weights change of a neuron is proportional to the learning rate parameter and gradient function. It basically involves two steps- forward and backward passes through multiple layers. The feature pattern is applied to the input in the forward direction and this input propagates by performing computations at each layer. The synaptic weights remain same in the forward pass. In the output layer the out value corresponding to the particular input is acquired and the error is computed between the resulting output and target value. This error is passed in backward direction with the computation of local gradient at each layer and changes the weights to get the minimum error for desired response. This forward and backward process run repeatedly till the goal is achieved [37], [30]**.**

3.3 Committee Neural Network (CNN)

Committee Neural Network (CNN) is another approach based on multiple and independent neural networks [37], [46], and [30]. The parallel structure of neural network resulting the final output by combining the outputs of its each member networks. This technique consists of 3 steps:

1. Selection of appropriate inputs for the individual member
2. Training of each member
3. Decision is based on majority basis

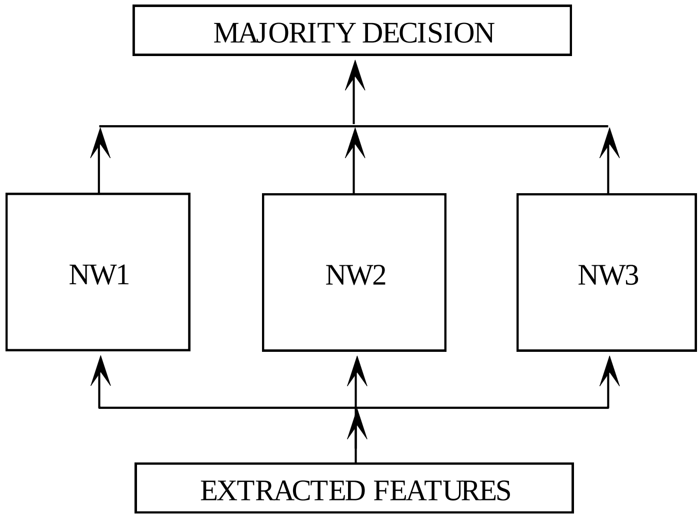


Fig. 3.2 Committee neural network [47]

CHAPTER 4
EXPERIMENTAL RESULTS

4.1 Normalization

This process can be implemented at any stage of processing. This limits the wide ranged varied signals to a small range variation. It is usually used before segment selection. The normalization is done by dividing each sample by the maximum absolute value present among the samples. It limits the signal in range of [-1, 1]. The normalized waveforms of three different EEG signals used for our study are shown in Fig. 4.1.

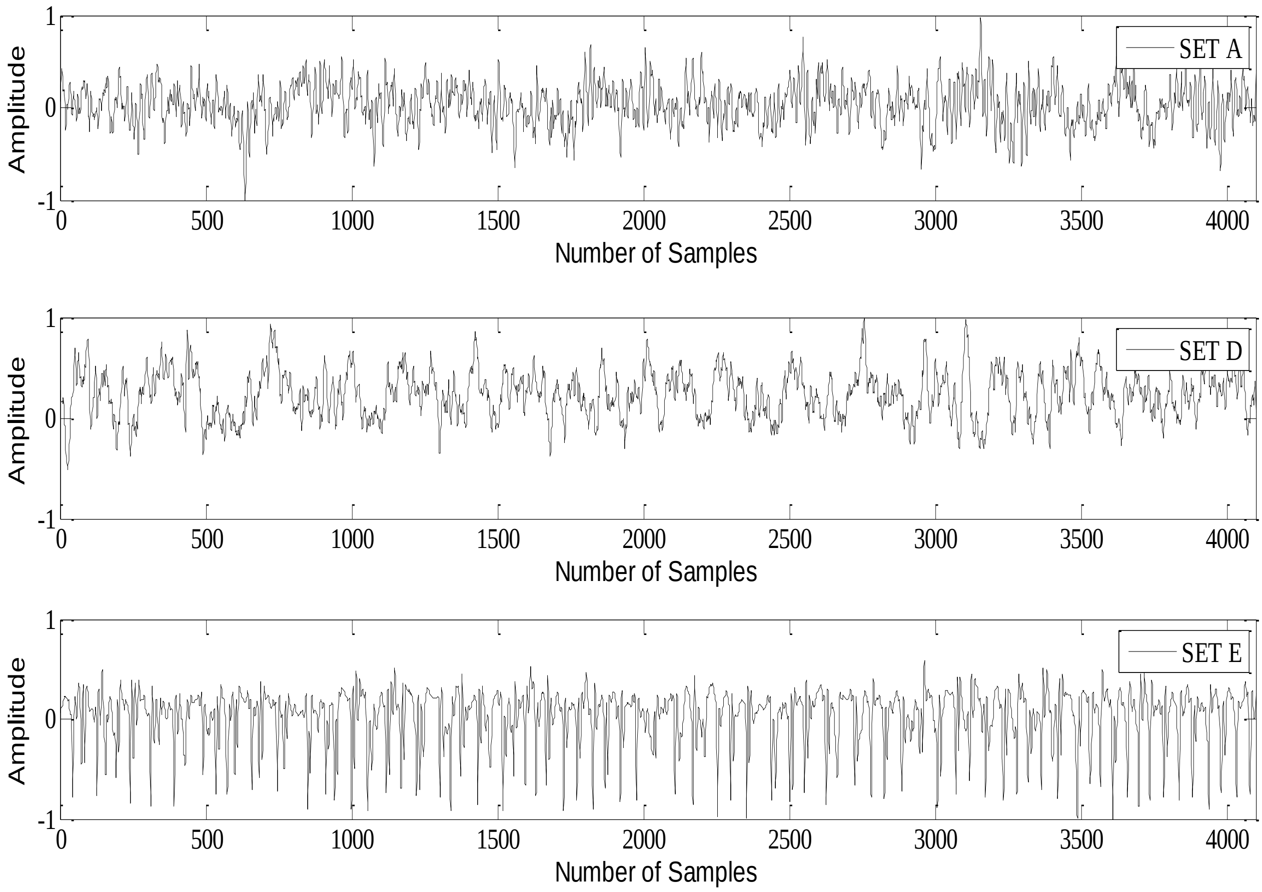


Fig. 4.1 Normalized waveform of SET A, SET D and SET E

4.2 Segment Selection

Each raw EEG signal is of 4097 samples which are very large for processing. The data is divided using windows of size 256 samples each to form the single EEG segment for processing. The total 16 EEG segments are obtained from single channel of raw signals. The single data set contains 100 channels of signals; so, the total 1600 EEG segments are obtained from single class. Hence, the total 4800 EEG segments (100 channels*16 EEG segments each) is obtained from three data sets [32]. The advantage of doing this faster processing is to obtain more samples for training and testing purposes. The Fig. 4.2 is shown with the plot of single EEG segment from each class.

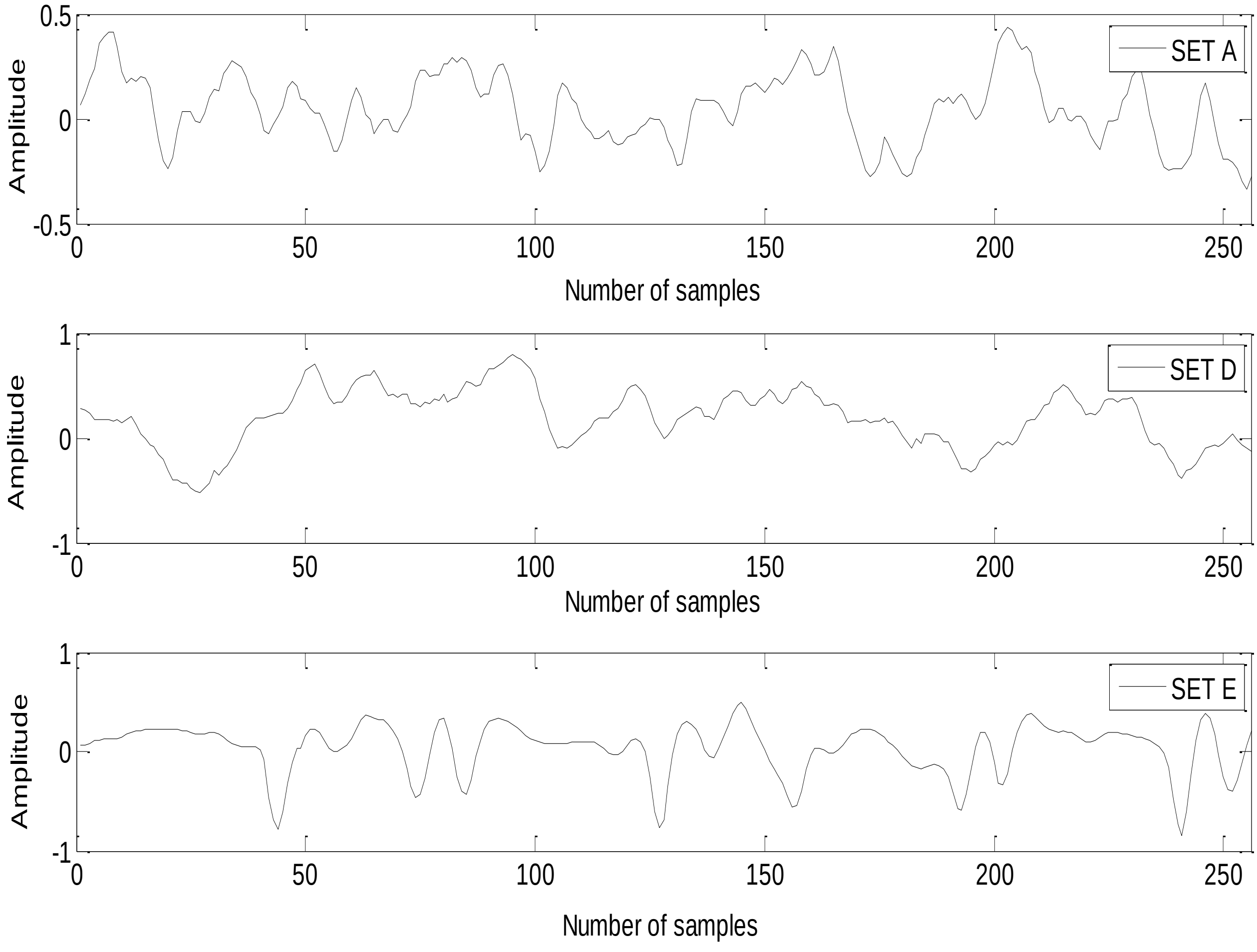


Fig. 4.2 Waveform of single EEG segment of Class A, Class D and Class E

4.3 Feature Extraction

The features are the most relevant and important characteristics of the signal. The extracted features are used for classification of original EEG signal by following DSP techniques. In the present study, the mainly three methods are used for feature extraction. The simulation graphs and results are given below for each step.

4.3.1 Analysis of Discrete Wavelet Coefficients

In the analysis of signal by using the discrete wavelet transform, the selection of mother wavelet and number of level decompositions is important. The highly popular wavelet type which is known as Daubechies Wavelet of order 2 is used for present study. We have chosen 4 levels decomposition. The smoothing nature of this mother wavelet enables it to detect the changes in fast varying signals like EEG signals. Each 256 length EEG segment is decomposed into four detail wavelet coefficients (d1, d2, d3, d4) and one approximation wavelet coefficients (a4). The notations d1, d2, d3, d4 or a4 represents the coefficients at first level, second level, third level and fourth level. The total 265 coefficients are obtained; among them, 247 are the detail coefficients (129, 66, 34, and 18) and 18 are the approximation coefficients [48]. The plots of the coefficients at each level of decomposition for single EEG segment of Class A, Class D and Class E is shown in Fig. 4.3.

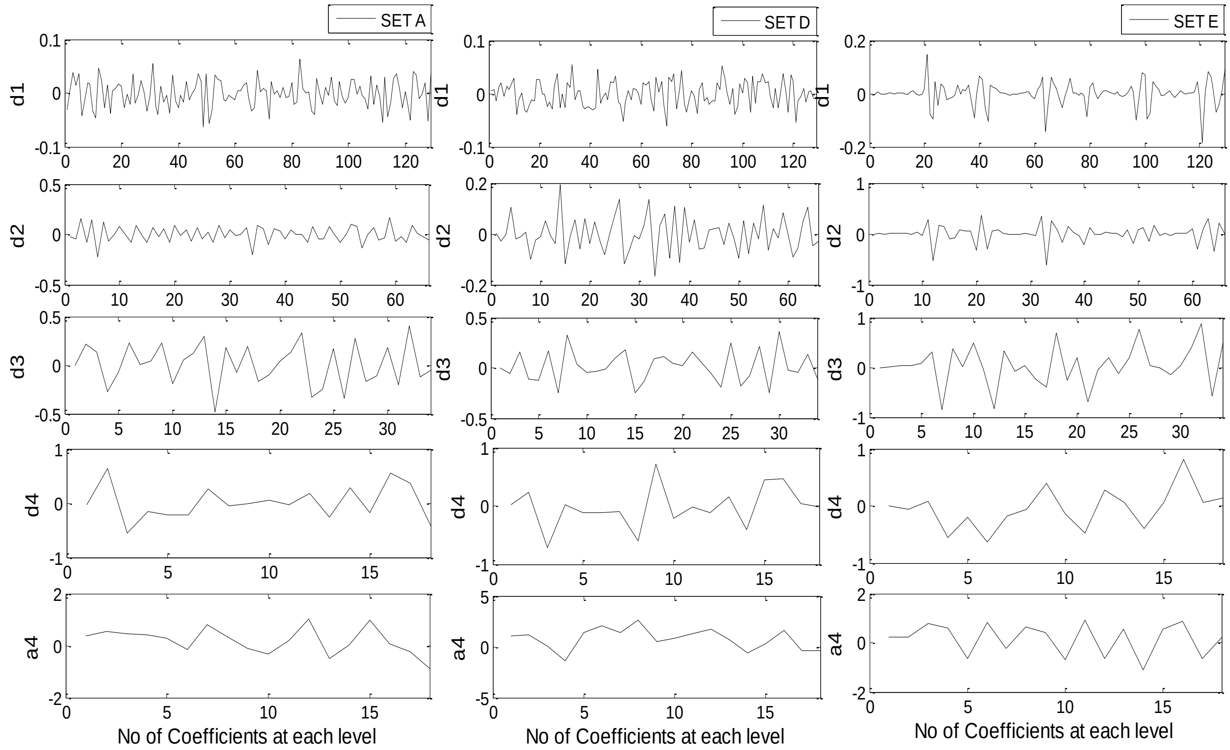


Fig. 4.3 Plots for wavelet and approximation coefficients of Class A, Class D, Class E of EEG signal at each decomposition level

The dimension of features is large so the following statistic is used to minimize the dimension.

i) *Maximum of wavelet coefficients in each sub band.*

ii) *Minimum of wavelet coefficients in each sub band.*

iii) *Mean of wavelet coefficients in each sub band*

iv) *Standard deviation of wavelet coefficients in each sub band*

Therefore, the 4 coefficients are obtained from each of the 5 sub bands resulting in a total of 20 coefficients feature vector. The plot of 20 dimension feature vectors for single EEG segment of Class A, Class D, and Class E. is shown in Fig. 4.4.TABLE 4.1 shows the features extracted from 3 different recordings.

TABLE 4.1 Features extracted from 3 different recordings of 3 classes

Data set	Extracted features	Sub-bands				
		D1	D2	D3	D4	A4
SET A	Maximum	0.0633	0.1647	0.3987	0.6316	1.01408
	Minimum	-0.0632	-0.2214	-0.4861	-0.5545	-0.9078
	Mean	-0.0013	0.0009	0.0084	0.0114	0.1811
	Standard deviation	0.0261	0.0781	0.2167	0.3176	0.5076
SET D	Maximum	0.05363	0.194	0.3605	0.7174	2.6052
	Minimum	-0.0596	-0.1667	-0.2514	-0.7248	-1.429
	Mean	-0.0007	-0.00015	0.0134	-0.0214	0.7655
	Standard deviation	0.02374	0.07099	0.1581	0.3547	1.0272
SET E	Maximum	0.1462	0.365	0.8636	0.8045	0.9287
	Minimum	-0.1843	-0.6088	-0.8549	-0.6272	-1.0864
	Mean	0.00E+00	5.00E-05	0.0371	-0.0437	0.1594
	Standard deviation	0.04	0.172	0.4057	0.348	0.645

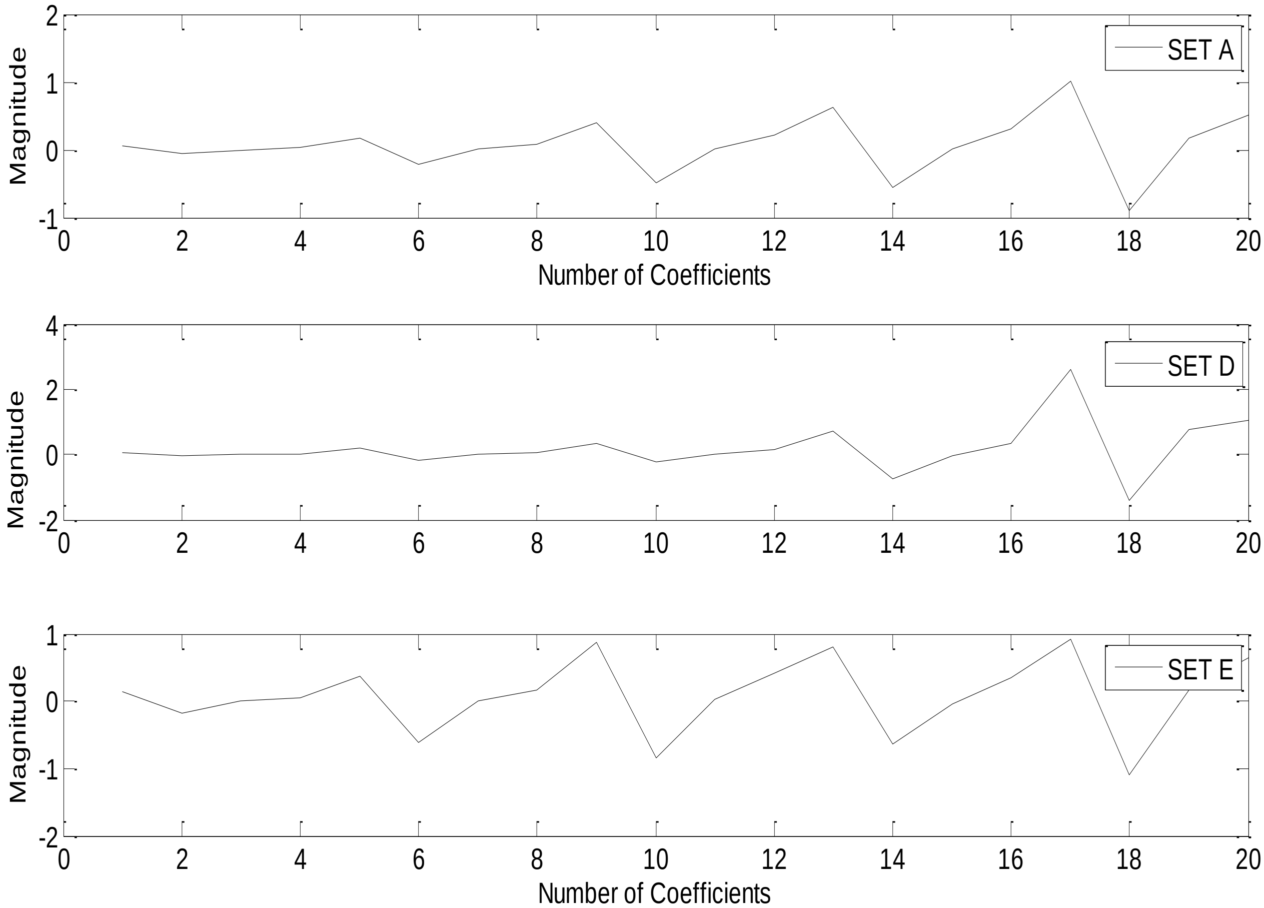


Fig. 4.4 Plots for 20 dimension Wavelet coefficients of Class A, Class D and Class E

4.3.2 Analysis of Autoregressive Coefficients

This method uses the Burg algorithm to fit the p^{th} order autoregressive (AR) model to the input signal by reducing the forward and backward errors while keeping the modeling parameters to satisfy the Levinson-Durbin recursion.

This estimates the reflection coefficient. Since, this is all pole model; so, the selection of optimal model order is required to preserve the stability. In the present study, the model order of 10 is chosen with reference from other literatures [29]. The 11 dimension feature vector is obtained for 10 model order. The power spectral density (PSD) by this method for the class A class D and class E is shown in Fig. 4.5.

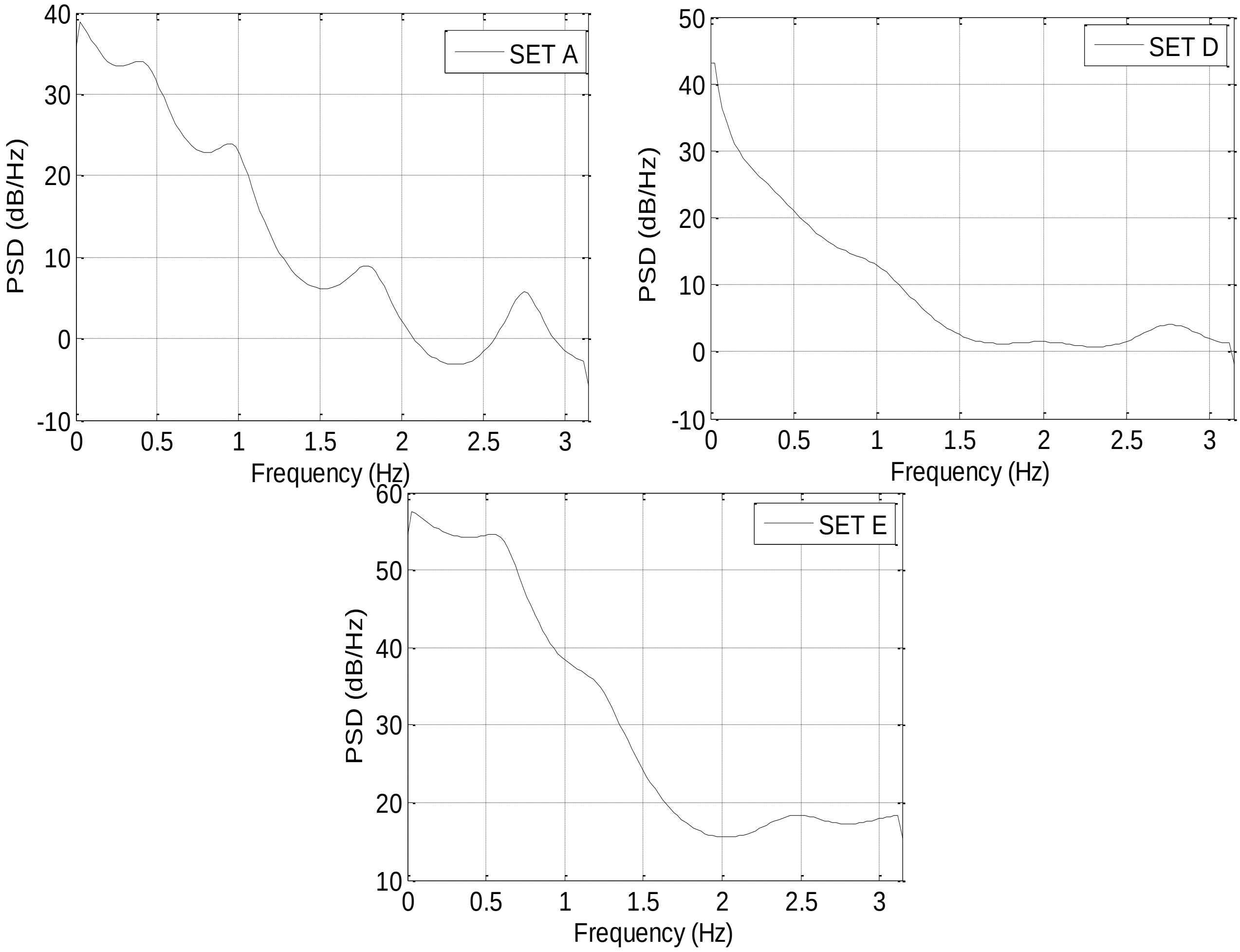


Fig. 4.5 Plots for power spectral density of Class A, Class D and Class E

4.3.3 Analysis of Lyapunov Exponents (LE)

Lyapunov exponents are the one of the widely used features which describes the qualitative nature of dynamical system. The following initials are chosen for finding the Lyapunov exponents in our work:

i) The length of embedding dimension (d) is equal to 128.

ii) The lag parameter (Δt)=1

iii) The evolution (m) is equal to 10.

iv) The upper limit of the neighboring point's lies inside the circle is 128.

The algorithm described in [42] is implemented for the analysis of each EEG segment and the 128 dimension feature vector is obtained for each EEG segment [42]. The Lyapunov exponents for three classes of EEG signals are shown in Fig. 4.6. The dimension of these features is large so the statistics method is used for reducing them as follows: TABLE 4.2 shows the features extracted from 3 different recordings.

i) *Maximum of Lyapunov exponents in each segment*

ii) *Minimum of Lyapunov exponents in each segment*

iii) *Mean of Lyapunov exponents in each segment*

iv) *Standard deviation of Lyapunov exponents in each segment.*

Thus the 4 dimension feature vector is obtained for each EEG segment.

TABLE 4.2 Features extracted from 3 different recordings of 3 classes

DATA SET	Extracted Features	Lyapunov Exponents
SET A	Maximum	0.29526405
	Minimum	0.00967177
	Mean	0.043100783
	Standard deviation	0.054876597
SET D	Maximum	0.303076263
	Minimum	0.015700801
	Mean	0.051319422
	Standard deviation	0.058299105
SET E	Maximum	0.319432519
	Minimum	0.008311364
	Mean	3.93E-02
	Standard deviation	0.057457055

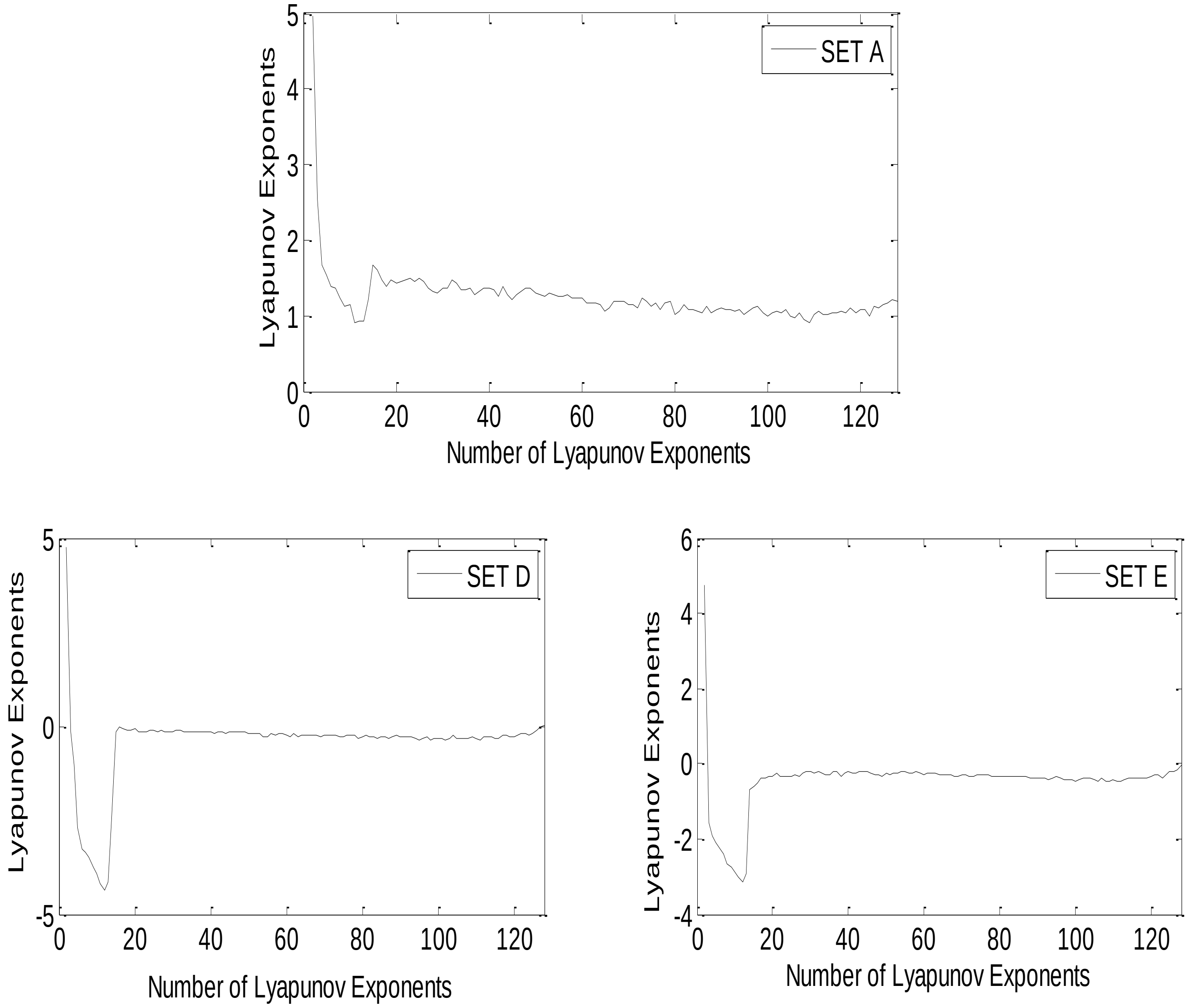


Fig. 4.6 Plots for Lyapunov exponents of Class A, Class D and Class E

4.3.4 Mixture of features

The feature vector is obtained for all EEG segments in a similar way as explained above for one EEG segment. Since these features are extracted by three methods. The purposed technique is based on the technique known as Mixture of features. The feature vectors obtained by two methods are appended for all EEG segments and the combining features are used as a feature vectors for training and testing purpose. In this report, the features are appended in two different combinations.

4.3.4.1 Mixture of DW coefficients and AR coefficients

After using statistics over features in DWT method, the length of feature vector obtained is equal to 20 for each segment and in case of autoregressive method the number of coefficients obtained is equal to 11 by selecting model order as 10 for each segment. After appending the feature vectors of DW coefficients and AR coefficients the 31 dimension feature vector is obtained. The appended feature vector in this combination is shown in Fig. 4.7. Similarly, we joined all EGG segments and used them for training and testing purposes.

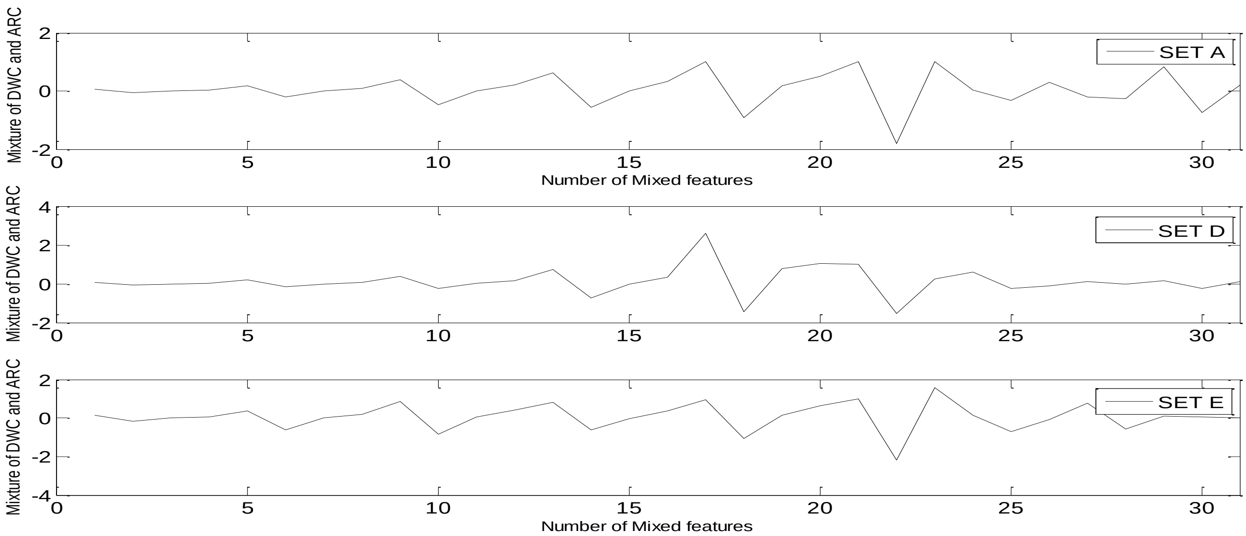


Fig. 4.7 Plots for mixing DW and AR coefficients of Class A, Class D and Class E

4.3.4.2 Mixture of DW coefficients, AR coefficients, and Lyapunov coefficients

In this method we use the features formed by appending DW, AR and Lyapunov coefficients. The 4 dimension feature vector is obtained after using statistics in Lyapunov exponents feature extraction based method. Thus the length of feature vector obtained by combining the feature extracted by three methods is 35(20 from DWT, 11 from AR and 4 from Lyapunov exponents). The plot for three different classes is shown in Fig. 4.8

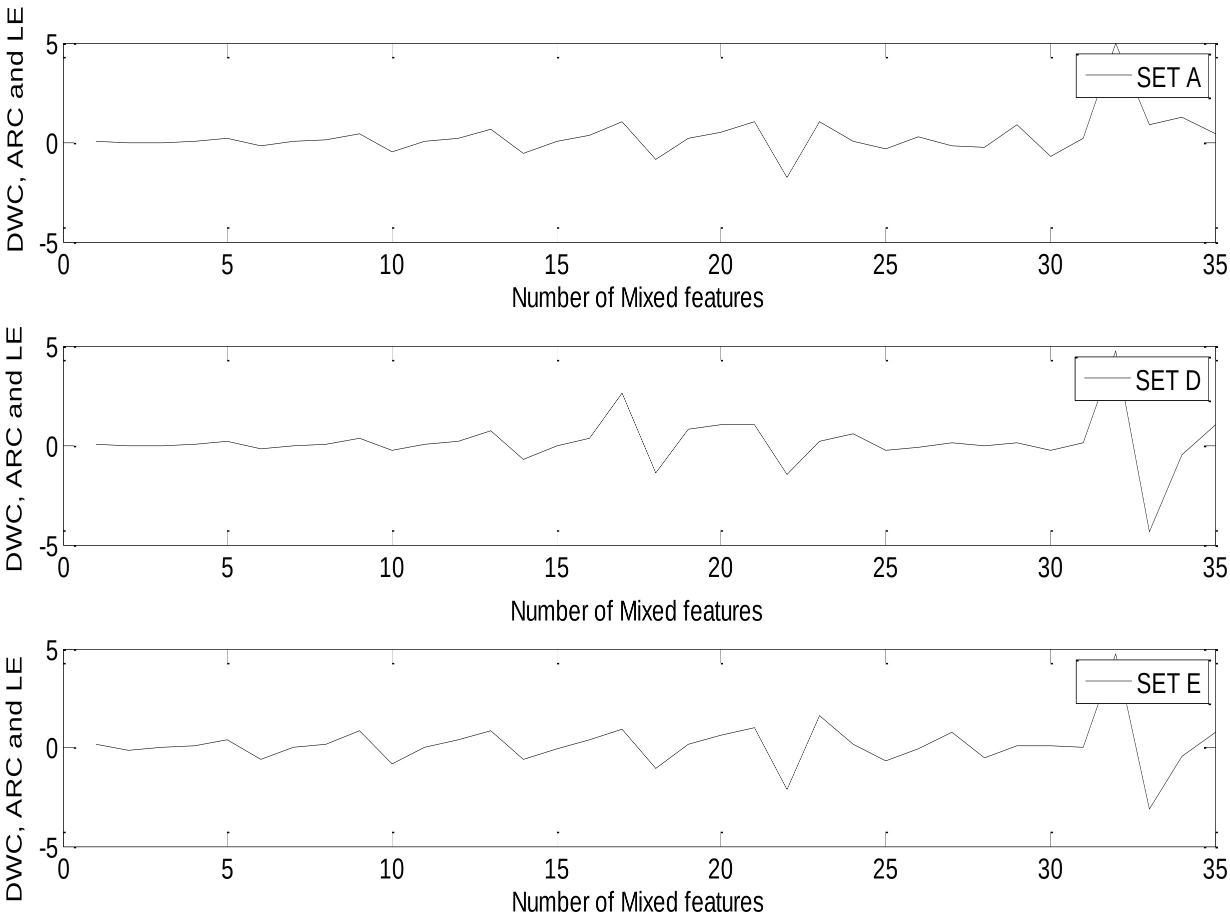


Fig 4.8 Plots for mixing DW, AR and LE coefficients of Class A, Class D and Class E

4.4 Implementation of different classifiers

4.4.1 Experiments for implementation of ANN

The classification is done by using MATLAB software package. The single MLPNN with Levenberg-Marquardt back-propagation algorithm is used to classify the EEG signals. The single hidden layer is chosen. In hidden layer and output, the sigmoidal activation function is used. The features computed by three methods are used for classification. In the present study, we have divided the features into two equal halves for training and testing. For the training and testing purpose we have used the extracted features separately as well as combined feature obtained after appending.

At the output, the target values used for generalize the network are

[0 0 1]= SET A

[0 1 0]= SET D

[1 0 0]= SET E

As we have already stated that the total 4800 EEG segments are formed from three classes of data (1600 segments from each class). The 2400 EEG segments are used for training (800 from each class) and 2400 segments are used for testing (800 from each class). These features are distinct for different classes. The features of all EEG segments are computed by using Discrete Wavelet Transform (DWT), Autoregressive (AR) model, Lyapunov exponents (LE) separately. The 20 dimension feature vector from DWT coefficients, 11 dimension feature vector from AR coefficient, 4 dimension feature vector from LE coefficients, 31 dimension feature vector obtained by combining DWT and AR coefficients and 35 dimension feature vector by combining DWT, AR & LE coefficients of each segment is supplied to the input of neural network one by one. The network is generalized and tested by supplying the testing features. The experimentation is performed many times with different hidden nodes to get the best results. The 800 samples of each class are tested and confusion matrices for majority decisions are made as shown in TABLE 4.3. The performance is calculated by implementing different features over the ANN classifier. The total classification accuracy obtained by our results is shown in TABLE 4.4.

4.4.2 Experiments for implementation of CNN

The committee is formed by 3 neural networks in our work. The decision is made on majority basis. In the CNN, the 2400 segments are used for training (800 from each class) and 2400 segments are used testing. The training data is divided among 3 networks.

i) The first neural network (NN1) is trained with 768 segments (256 from each class).

ii) The second neural network (NN2) is trained with 768 segments (256 from each class).

iii) The third neural network is trained with 288 segments (288 from each class)

The each MLPNN with Levenberg-Marquardt back-propagation algorithm is used for classify the EEG signals. The single hidden layer is chosen for all the members. The each network is tested by training with different hidden nodes for considering the best results. The confusion matrices based on majority decision is formed and shown in TABLE 4.3. The performance is calculated by implementing different features over the CNN classifier. The total classification accuracy is shown in TABLE 4.4.

TABLE 4.3 Confusion matrices of ANN, NN1, NN2, NN3 and CNN

Classifier	Desired result	Output results														
		DWT method			AR method			LE method			DWT+AR			DWT+AR+LE		
		Set A	Set D	Set E	Set A	Set D	Set E	Set A	Set D	Set E	Set A	Set D	Set E	Set A	Set D	Set E
ANN	Set A	725	20	55	795	5	0	673	15	112	789	28	8	781	7	12
	Set D	63	688	49	21	732	47	48	674	78	11	768	17	45	716	39
	Set E	54	59	687	6	14	770	159	87	554	0	34	775	21	24	755
NN1	Set A	678	36	86	784	14	2	475	8	117	767	24	9	779	14	7
	Set D	107	607	54	50	701	49	79	656	65	34	730	36	57	716	27
	Set E	86	78	668	21	102	677	209	74	517	26	64	710	29	27	726
NN2	Set A	652	43	105	744	53	3	610	10	180	746	48	6	761	23	16
	Set D	78	666	56	25	723	52	56	694	50	26	746	33	35	732	33
	Set E	48	138	614	9	44	747	153	121	526	13	33	754	12	30	758
NN3	Set A	680	45	75	775	21	4	620	25	105	781	10	9	782	5	13
	Set D	52	624	124	29	654	117	34	691	75	41	678	81	53	663	84
	Set E	76	60	658	12	6	782	159	103	538	21	4	775	17	3	780
CNN	Set A	707	21	72	787	8	5	685	10	105	788	9	3	779	11	10
	Set D	61	658	81	23	711	66	57	678	65	18	745	37	31	725	44
	Set E	41	58	701	12	16	772	182	87	531	16	12	772	17	19	764

TABLE 4.4 Total classification accuracy of ANN, NN1, NN2, NN3 and CNN

Classifier	Accuracy (in %)				
	DWT	AR	LE	DWT+AR	DWT+AR+LE
ANN	85.50	94.91	79.21	95.91	93.83
NN1	81.37	90.08	77	91.96	92.54
NN2	80.50	92.25	76.25	93.58	93.79
NN3	81.75	92.12	79.12	93.08	92.71
CNN	86.08	94.58	78.92	96.04	94.50

CHAPTER 5
CONCLUSION & FUTURE WORK

Conclusion

The EEG signal was collected from the standard data base. These EEG signals were not distinguishable with human eyes. We used the signal processing tools to distinct them and provide the status of the individual.

- It was found that the feature extraction by Autoregressive model using Burg algorithm gives the best performance as compared to DWT and Lyapunov Exponents algorithms.
- For the classification schemes, it was found that the CNN algorithm clearly outperforms over ANN algorithm.
- The purposed technique successfully used the combined features for EEG signals. The performance was found to be best using combined features extracted by DWT and Autoregressive model. It was observed that the performance decreased by adding the features extracted from Lyapunov Exponents because the performance using only LE as feature extraction method was found to be least among the three feature extraction methods.

Future Work

- We can use other algorithms and techniques for the feature extraction and classification of EEG signals to further improve the accuracy of the identification system.
- We can further improve the system by reducing the complexity. The main objective could be to find the best algorithms which optimize the performance and complexity.
- We can use the implemented algorithms for other applications of EEG signals as well as for other biometric identification systems.

REFERENCES

[1] E. Niedermeyer and F.L. da Silva, "Electroencephalography: Basic Principles, Clinical Applications, and Related Fields," Lippincot Williams & Wilkins, 2004.

[2] http://en.wikipedia.org/wiki/Electroencephalography.

[3] B. Abou-Khalil and K.E. Musilus, "*Atlas of EEG & Seizure Semiology*," Elsevier Health Sciences, 2006.

[4] Saeid Sanei and J.A. Chambers, *EEG Signal Processing*, John Wiley & Sons Ltd, England, 2007.

[5] R. Caton, "The electric currents of the brain," *Br. Med. J.*, vol. 2, pp. 278, 1875.

[6] W.G. Walter, "Slow potential waves in the human brain associated with expectancy, attention and decision," *Arch. Psychiat. Nervenkr.*, vol. 206, pp. 309 – 322, 1964.

[7] A. Massimo, "In Memoriam Pierre Gloor (1923 – 2003): an appreciation," *Epilepsia*, vol. 45, no. 7, pp. 882, July 2004

[8] H.L. Attwood and W. A. MacKay, *Essentials of Neurophysiology*, B. C. Decker, Hamilton, Canada, 1989.

[9] M. Teplan, "Fundamentals of EEG measurements," *Measmt Sci. Rev.*, vol. 2, no. 2, 2002.

[10] R.D. Bickford, *Electroencephalography, in Encyclopedia of Neuroscience*, Ed. G. Adelman, Birkhauser, Cambridge (USA), pp. 371 – 373, 1987.

[11] H. Jasper, "Report of committee on methods of clinical exam in EEG," *Electroencephalogr. Clin. Neurophysiol.*, 10, pp. 370 – 375, 1958.

[12] E. Niedermeyer, *The normal EEG of the waking adult, in Electroencephalography, Basic Principles, Clinical Applications, and Related Field*, Eds E. Niedermeyer and F. Lopes da Silva, 4th Edition, Lippincott, Williams and Wilkins, Philadelphia, Pennsylvania, pp. 174 – 188, 1999.

[13] G. Pfurtscheller, D. Flotzinger, and C. Neuper, "Differentiation between finger, toe and tongue, movement in man based on 40 Hz EEG," *Electroencephalogr. Clin. Neurophysiol.,* vol.90, pp. 456 – 460, 1994.

[14] S. Ashwal and R. Rust, "Child neurology in the 20th century," *Pedia. Res.,* vol. 53, pp. 345 – 361, 2003.

[15] E.D. Adrian, and B.H.C. Mattews, "The Berger rhythm, potential changes from the occipital lob in man," *Brain*, vol. 57, pp. 345 – 359, 1934.

[16] J. Trabka, "High frequency components in brain waves," *Electroencephalogr. Clin. Neurophysiol.*, vol.14, pp. 453 – 464, 1963.

[17] http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html.

[18] E.D. Übeyli, "Lyapunov exponents/probabilistic neural networks for analysis of EEG signals," *Expert Systems with Applications*, vol. 37, no. 2 pp. 985-992, 2010.

[19] http://ethesis.nitrkl.ac.in/1818/1/thesis_9021_9011.pdf.

[20] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David and C.E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state," *Physical Review E*, vol. 64, 1998.

[21] http://www.dtic.upf.edu/~xserra/cursos/TDP/referencies/Park-DWT.pdf.

[22] M. Unser, A. Aldroubi, "A review of wavelets in biomedical applications," *Proceedings of the IEEE,* vol. 84, no. 4, pp. 626–638, 1996.

[23] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory,* vol.36, no. 5, pp. 961–1005, 1990.

[24] S. Soltani, "On the use of the wavelet decomposition for time series prediction," *Neurocomputing*, vol. 48, pp. 267–277, 2002.

[25] İ. Güler and E.D. Übeyli, “Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients,” *Journal of Neuroscience Methods*, vol.148, no. 2, pp. 113-121, Oct. 2005.

[26] J.G. Proakis, D.G. Manolakis, *Digital Signal Processing*, Pearson Education, Pearson, 4th Edition.

[27] J.P. Burg, “Maximum entropy spectral analysis,” *in Proc. 37th Meeting Soc. Exploration Geophys.*, 1967

[28] P. Stoica and R.L. Moses, *Introduction to spectral Analysis*, Upper Saddle River, NJ: Prentice-Hall, 1997.

[29] E.D. Übeyli, “Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals,” *Expert Systems with Applications*, vol. 37, pp. 233–239, 2010.

[30] http://ethesis.nitrkl.ac.in/1818/1/thesis_9021_9011.pdf.

[31] R. Silipo, G. Deco, R. Vergassola, and H. Bartsch, “Dynamics extraction in multivariate biomedical time series,” *Biological Cybernetics*, vol. 79, no. 1, pp. 15–27, 1998.

[32] N.F. Güler, E.D. Übeyli, and I. Güler, “Recurrent neural networks employing Lyapunov exponents for EEG signals classification,” *Expert Systems with Applications*, vol. 29, no. 3, pp. 506–514, 2005.

[33] A.M. López Jiménez, C. Camacho Martínez Vara de Rey, and A.R. García Torres., “Effect of Parameter Calculation in Direct Estimation of the Lyapunov Exponent in Short Time Series,” *Discrete Dynamics in Nature and Society*, vol. 7, no. 1, pp. 41-52, 2002.

[34] H.D.I. Abarbanel, R. Brown, and M.B. Kennel, "Lyapunov exponents in chaotic systems: Their importance and their evaluation using observed data," *International Journal of Modern Physics B*, vol. 5, no. 9, 1991.

[35] M. Casdagli, "Nonlinear prediction of chaotic time series," *Physica D*, vol. 35, no. 3, pp. 335–356, 1989.

[36] J.P. Eckmann, S.O. Kamphorst, D. Ruelle, and S. Ciliberto, "Lyapunov exponents from time series," *Physical Review A*, vol. 34, no. 6, 4971–4979, 1986.

[37] S. Haykin, X.B. Li, "Detection of signals in chaos," *Proc. IEEE*, vol. 83, no.1, pp. 95–22, 1995.

[38] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, and L.S. Tsimring, "The analysis of observed chaotic data in physical systems," *Reviews of Modern Physics*, vol. 65, no. 4, pp. 1331–1392, 1993.

[39] G. Benettin, L. Galgani, A. Giorgilli, and J.M. Strelcyn, "Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them," *Meccanica*, vol. 15, no. 1, pp. 9–30, 1980.

[40] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.

[41] A. Wolf, J.B. Swift, H.L. Swinney, and J.A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D*, vol.16, no. 3, pp. 285–317, 1985.

[42] M. Sano, and Y. Sawada, "Measurement of the Lyapunov spectrum from a chaotic time series," *Physical Review Letters*, vol. 55, no. 10, pp. 1082–1085, 1985.

[43] E.D. Übeyli, "Analysis of EEG signals using Lyapunov exponents," *Neural Network World*, vol.1, no. 3, pp. 257–273, 2006.

[44] E.D. Übeyli, “Fuzzy similarity index employing Lyapunov exponents for discrimination of EEG signals,” *Neural Network World,* vol.16, no. 5, pp. 421–431, 2006.

[45] M.T. Rosenstein, J.J. Collins, and C.J. De Luca “A practical method for calculating largest Lyapunov exponents from small data sets,” *NeuroMuscular Research Center and Department of Biomedical Engineering*, Boston University, 44 Cummington Street, Boston, MA 02215, US.

[46] A.J.C. Sharkey, “On combining artificial neural nets,” *Connect. Sci.*, vol. 8, pp. 299–314, 1996.

[47] N.P. Reddy and O.A. Buch, “Speaker verification using committee neural networks”, *Computer Methods and Programs in Biomedicine*, vol. 72, no. 2, pp.109-115, 2003.

[48] İ. Güler and E.D. Übeyli, “Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients,” *Journal of Neuroscience Methods*, vol. 148, no. 2, pp. 113-121, Oct. 2005.