EEG SIGNAL CLASSIFICATION USING LDA AND MLP CLASSIFIER

Romain Atangana[1,2,3], Daniel Tchiotsop[1], Godpromesse Kenne[1], Laurent Chanel DjoufackNkengfac k[1,2],

[1]Unité de Recherche d'Automatique et d'InformatiqueAppliquée (UR-AIA), IUT-FV of Bandjoun, University of Dschang-Cameroon, P.O. Box 134-Bandjoun
[2]Unité de Recherche de Matière Condensée-d'Électronique et de Traitement du Signal (UR-MACETS), Faculty of Science, University of Dschang-Cameroon, P.O. Box 67-Dschang
[3]Division of Continuing Education and Distance Learning, Higher Teacher Training College, University of Ngaoundere-Cameroon, P.O. Box 652-Bertoua

ABSTRACT

Electroencephalography (EEG) is the recording of electrical activities along the scalp. EEG measures voltage fluctuations resulting from ionic current flows within the neurons of the brain. Diagnostic applications generally focus on the spectral content of EEG, which is the type of neural oscillations that can be observed in EEG signal. EEG is most often used to diagnose epilepsy, which causes obvious abnormalities in EEG readings. This powerful property confirms the rich potential for EEG analysis and motivates the need for advanced signal processing techniques to aid clinicians in their interpretations. This paper describes the application of Wavelet Transform (WT) for the processing of Electroencephalogram (EEG) signals. Furthermore, the linear discriminant analysis (LDA) is applied for feature selection and dimensionality reduction where the informative and discriminative two-dimension features are used as a benchmark for classification purposes through a Multi-Layers Perceptron (MLP) neural network. For five classification problems, the proposed model achieves a high sensitivity, specificity and accuracy of 100%.Finally, the comparison of the results obtained with the proposed methods and those obtained with previous literature methods shows the superiority of our approach for EEG signals classification and automated diagnosis.

KEYWORDS

EEG, Discrete Wavelet Transform (DWT), Linear Discriminant Analysis (LDA), Multi-Layers Perceptron (MLP)Classifier

1. INTRODUCTION

In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a short period of time, usually 20 to 40 minutes as recorded from multiple electrodes placed on the scalp. The brain's electrical charge is maintained by billions of neurons. Neurons are electrically charged (or "polarized') by membrane transport proteins that pump ions across their membranes. Neurons are constantly exchanging ions with the extracellular milieu for example to maintain resting potentials and to propagate action potentials. Ions of similar charge repel each other, and when many ions are pushed out of many neurons at the same time, they can push their neighbors, who push their neighbors, and so on, in a wave. This process is known as volume conduction. When the wave of ions reaches the electrodes on the scalp, they can push electrons on the metal on the electrodes. Since metal conducts the push and pull of electrons easily, the difference in push or pull voltages between any two electrodes can be measured by a voltmeter. Recording these voltages over time gives us the EEG.

The electric potential generated by an individual neuron is far too small to be picked up by EEG. EEG activity therefore always reflects the summation of the synchronous activity of thousands or millions of neurons that have similar spatial orientation. If the cells do not have similar spatial orientation, their ions do not line up and create waves to be detected. Pyramidal neurons of the cortex are thought to produce the most EEG signal because they are well aligned and fire together. Scalp EEG activity shows oscillations at a variety of frequencies. Several of these oscillations have characteristics frequency ranges, spatial distributions and are associated with different states of the brain functioning. These oscillations are referred to as sub-bands based on the frequency ranges. These bands from low to high frequencies respectively are called delta(δ) (range 0.5 - 4Hz), theta (θ) (range 4-8Hz), alpha (α) (range 8- 13Hz), beta (β) (range 13-32Hz), and gamma (γ) (range 32-45Hz) [1].

The visual distinction of seizure from common artifacts within an EEG measurement is based on the shape and spikiness of the waveforms. A signal with seizure has a rhythmical and prominent spike, whereas most of other artifacts are non-stationary and randomly shaped. But considering the fact that the recorded EEG pattern is a special mapping of signals captured by placement of electrodes onto different regions on the scalp, it is extremely difficult for human beings to observe and understand the actual behavior of the brain by merely visual inspection. Hence there is an ever-increasing demand of an easily accessible and fully automated epileptic seizure detection system using EEG signals.

During this last decade, advancements in biomedical signal processing have led to the analysis of EEG being increasingly numerical for automated diagnosis of brain diseases. Epilepsy is the most prevalent neurological disorder in humans and is characterized by recurrent seizures [2]. Seizures are defined as sudden changes in the electrical functioning of the brain, resulting in altered behaviors, such as losing consciousness, jerky movements, temporarily loss of breath and memory loss. Most of epilepsy analysis methods developed in the 20th century were based on the concept of visual inspection of EEG signals by skilled electroencephalographers. However, with the advancement in the field of signal processing and pattern recognition, different automatic techniques of epileptic seizure have been developed in the last two decades [3-4].

Although the occurrence of epileptic seizures seems unpredictable [3], exciting application of seizure prediction technology in therapeutic treatment of epilepsy can be used to trigger intervention in preventing seizures before they begin [5].

Recently, seizure prediction was extremely investigated through different models of automatic EEG signals classification and seizure detection [6-24]. These models generally include three steps and widely vary in their theoretical approaches of the problem, validation of results and the amount of data analyzed. In early attempts to show a relationship between the raw EEG and behavior, the first step is generally related to the decomposition of the raw data and the feature extraction processes. The decomposition process is applied in order to explore the raw data in frequency domain [19,23], in time-frequency domain [3, 10,15,16,18,22], and in multi-domain [8,12,19,23]. After the decomposition process, statistic features and entropies features [6, 10,14,[20] were widely extracted. Since these extracted features are generally in high dimension and are no more discriminative for different classes, the second step proposed a selection of an optimal and low dimension features through a dimensionality reduction process. For this purpose, principal component analysis (PCA), independent component analysis (ICA), LDA and K-means clustering were applied [6]. Then, at the last step, these most prominent and discriminative features are used as inputs of a kernel machine for the classification decision. For this purpose, two mostly used kernel machines with different architectures were proposed, such as artificial neural network (ANN) [3,7,9,11,13,21] and support vector machine (SVM) [6,14,17,22]. Most of these studies focus on epilepsy by using a feature vector with at least three dimensions as inputs of a complex kernel machine.

This paper applies WT to explore EEG signals in the time-frequency domain. EEG data are projected into a set of wavelet coefficients that exhibit the correlation which exist between the signals and their corresponding EEG rhythms. Furthermore, Areas under the frequency spectrum of each wavelet coefficient are used as the features. These areas are calculated using trapezoidal rule from numerical analysis. For feature extraction purposes we first transform each wavelet coefficient into its time domain then we apply Fourier transform to the wavelet coefficient to obtain the frequency spectrum [25, 26] LDA analysis is mainly used to classify two class problems. The main task of LDA is to use hyper plane to separate or to discriminate data. LDA is a predictable discriminant analysis method which enables us to predict the membership of a sample of a class that has been predefined by a set of features. On a practical point of view, this method finds a projection of original data on a more reduced space so that features of different classes should be well separated.

The rest of this paper is organized as follows: the second section describes the EEG dataset used. The EEG analysis using WT technique for rhythms separation is briefly presented. LDA and MLP are also presented for the purpose of feature extraction and classification, respectively. Simulation results are presented and discussed in section 3. Then, the last part summarizes this work.

2. MATERIALS AND METHODS

2.1. Dataset description

In the last few years, most researchers used publicly available data described in [28] for their research work in the field of epileptic seizure detection. We are also using the same benchmark database in order to compare our results with the results of previous research works. The database is prepared by taking inputs from different subjects and is divided into five Sets (A – E) each containing 100 EEG samples recorded through single channel. The mental state of the subjects in each data set at the time of data recording was as follows:

- A - B: Five healthy volunteers relaxed in an awake state with eyes open (A) and closed (B);
- C- D: Activities measured during seizure free periods of EEG from five patients, all of whom had achieved complete seizure control and were correctly diagnosed;
- E: Contains seizure activity.

The dataset was recorded using a 128-channel amplifier system and standardized 10-20 electrode placement scheme. After recording, the data were sampled and digitized at 173.61 Hz in a time duration of 23.6 seconds using 12 bits resolution as shown in table 1. Therefore, the frequency range with the characteristics mentioned above is from 0.05 – 86.80 Hz. For extracting the feature from EEG samples, the frequency range of interest is 0-64 Hz.

The dataset used in this study was collected from ten patients with a sampling frequency of 173.61 Hz and consists of five sets as presented in Table 1.

Table 1: EEG dataset description [28].

Set	Number of Segments	Record Mode	Record Type	Number of Samples / Duration (s)
SET A	100	Extracranial	Normal (eyes opened)	4097 / 23.6

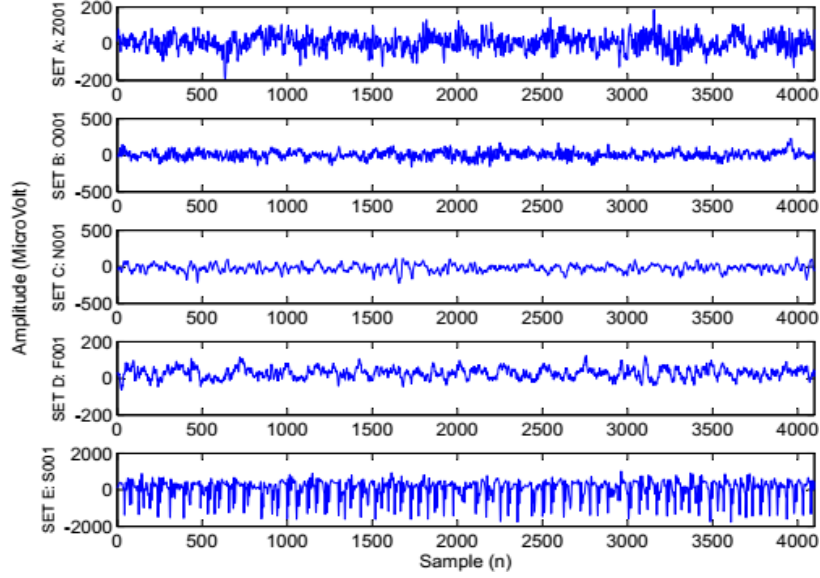| SET B | 100 | Extracranial | Normal (eyes closed) | 4097 / 23.6 |
SET C	100	Intracranial	Pre-ictal	4097 / 23.6
SET D	100	Intracranial	Pre-ictal	4097 / 23.6
SET E	100	Intracranial	Epileptic	4097 / 23.6

Figure 1: Examples of EEG signals.

2.2. Wavelet transform (WT) analysis

EEG signals may contain non-stationary characteristics. Thus, efficient and instantaneous separation of the EEG rhythms can be done based on the decomposition of the EEG signals into wavelet coefficients.

Wavelet transform (WT) is a powerful spectral estimation technique for the time-frequency analysis of a signal [29]. The continuous WT (CWT) of a signal $x(t)$ is defined as:

$$CWT_{a,b}(x) = \frac{1}{\sqrt{|a|}} \int x(t)\psi^*\left(\frac{t-b}{a}\right)dt \qquad (1)$$

Where $\psi(t)$ is called the mother wavelet function (MWT), while a and b ($a,b \in IR$) are dilation and translation parameter, respectively.

The application of WT in engineering areas usually requires the discrete WT (DWT). DWT, which is non-redundant and has less computational time than CWT. DWT is defined by using the discrete values of the parameters a and b.

$$DWT_{m,n}(x) = a_0^{-m/2} \int x(t)\psi(a_0^{-m/2}t - nb_0)dt \qquad (2)$$

Where m and n ($m,n \in Z$) are indicating frequency and time localization, respectively.

For a dyadic-orthonormal WT, a_0 and b_0 are set to 2 and 1, respectively. This choice provides the basis for multi-resolution analysis. For the purpose, Mallat[30] developed a dyadic method

by which DWT separates a signal into low and high-frequency components via two finite impulse response quadrature mirror filters, where g a high-pass filter is related to the MWT and h is a low-pass filter related to the scaling function.

$$g(h) = (-1)^n h(1-n) \tag{3}$$

$$\phi(x) = \sum_n h(n)\sqrt{2}\phi(2x-n) \tag{4}$$

$$\psi(x) = \sum_n g(n)\sqrt{2}\phi(2x-n) \tag{5}$$

The result transforms the original signal into two equal sub-bands $[0 - F_{N/2}]$ and $[F_{N/2} - F_N]$ and the high-frequency component can be sent to the second stage to repeat the procedure as presented by Figure 2. Finally, the signal is decomposed at the expected level. The MWT and the wavelet decomposition level should be selected carefully.
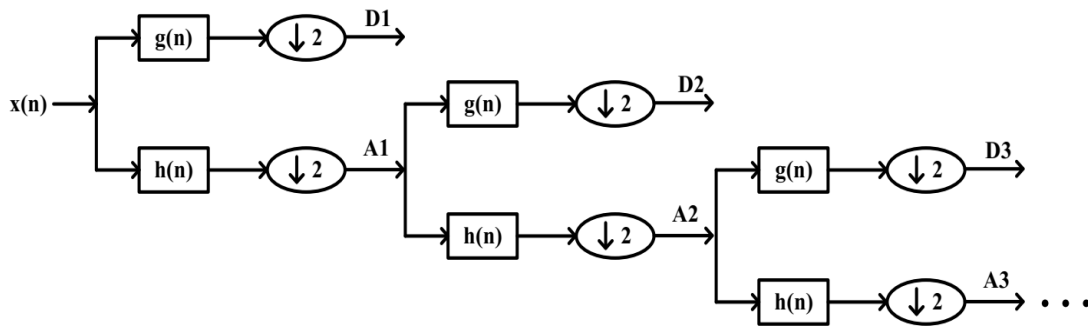

Figure 2: Sub-bands separation using WT.

Table 2.Rhythms, corresponding levels and simulated frequencies for a four-level decomposition using Db4 MWT.

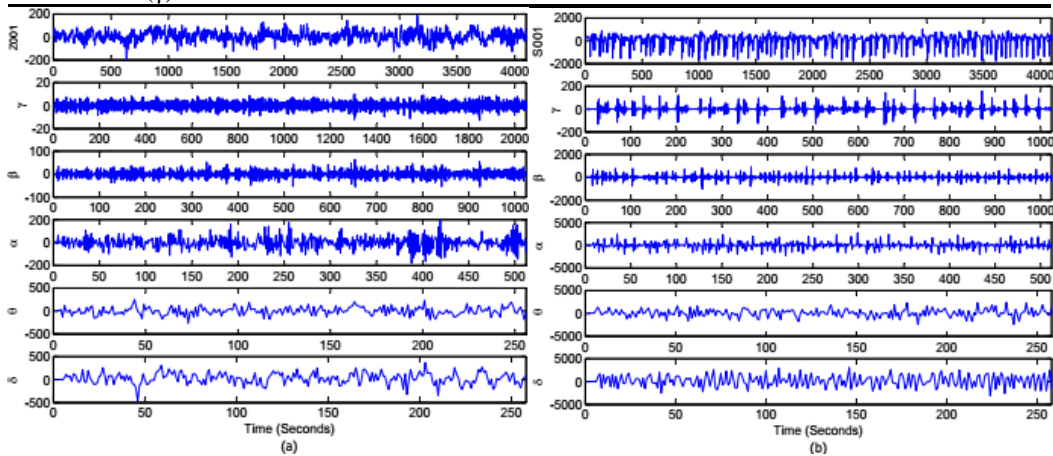| EEG rhythm	Separated signal	Frequency range (Hz)	Simulated frequency (Hz)	
			Set A EEG	Set E EEG
Delta (δ)	A4	0-4	0.21	3.60
Theta (θ)	D4	4-8	5.76	5.97
Alpha (α)	D3	8-16	11.53	14.11
Beta (β)	D2	16-32	26.10	26.19
Gamma (γ)	D1	32-64	50.00	64.33

Figure 3: Wavelet coefficients for the EEG rhythms separation of (a) a normal EEG Z001 and (b) an ictal EEG S001.

After rhythms separation with the WT, the frequency distribution of the EEG signal is analyzed by calculation of areas under the frequency spectrum of each wavelet coefficient through the trapezoidal rule. Trapezoidal rule aims to approximate integral $\int_a^b f(x)\, dx$ by the sum of the "areas" of the trapezoids as shown in Figure 4.
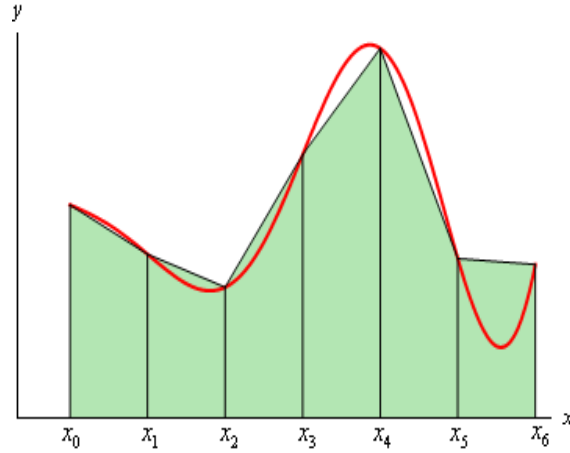

Figure 4: Trapezoidal approximation to $\int_a^b f(x)\, dx$

$$\int_a^b f(x)dx = \frac{1}{2}(y_0 + y_1)h + \frac{1}{2}(y_1 + y_2)h + \cdots + \frac{1}{2}(y_{n-1} + y_n)h \qquad (5)$$

In which $y_j = f(x_j) for the choosen partition a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$.

And h is the spacing between successive points. If we set, $h = \frac{(b-a)}{n}$ then the trapezoidal rule can be written as:

$$\int_a^b f(x)dx = \frac{(b-a)}{2n}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n) \qquad (6)$$

2.3. Linear discriminant analysis (LDA) feature Selection

Once that we have our feature vector, in order to increase the performance of the classifier it is advisable to reduce its dimensionality by keeping the representative information in the feature vector and eliminating the redundant one from it. The resulting vector is called reduced feature vector. There are two main strategies of dimensionality reduction.

- Feature projection: This strategy tries to determine the best combination of the original features to form a new feature set, generally smaller than the original one [31]. Principal Component Analysis (PCA) is using this type of strategy, it is an orthogonal linear transformation that can be used for dimensionality reduction in a signal database by retaining those characteristics of the dataset that contribute most to its variance in order to do this; lower order principal components are kept and higher order ones are dismissed [32].This could be applied to EEG signals.

- Feature Selection: This strategy chooses the best subset of the original feature vector according to some criteria for judging whether one subset is better than another. The ideal criterion for classification should be the minimization of the probability of misclassification, but generally simpler criteria based on class separability are chosen [31] Linear Discriminant Analysis (LDA) is used in that type of strategy.

LDA analyzes signals to extract the most prominent features. The main aim of LDA feature extraction is to obtain further information from the signal by transforming the feature vector formed by the areas into fewer ones as a feature vector. The extracted features are more informative and discriminative with reducednoise and redundancy compared to the original data.

LDA is a supervised algorithm since it computes the linear discriminant features by maximizing distance between classes and minimizing distance within classes [6]. The LDA projection can be written as a single matrix equation described by:

$$Y = X * W \tag{7}$$

where Y and X are the original data and features, respectively. And W is the *dxk* projection matrix formed of the first *k*-eigenvectors of $S_w^{-1} S_b$, where S_w and S_b are the within-class and between-class scatter matrices, respectively.

$$S_w = \frac{1}{N} \sum_{l=1}^{L} \sum_{i=1}^{N_l} \left(x_i^l - \mu_l \right)^T \left(x_i^l - \mu_l \right) \tag{8}$$

$$S_b = \frac{1}{N} \sum_{l=1}^{L} N_l \left(\mu_l - \mu \right)^T \left(\mu_l - \mu \right) \tag{9}$$

x_i^l, N_l and μ_l are an i-row vector, the number of i-vectors and the mean vector corresponding to class label l, respectively. L, N and μ are the number of classes, the total number of i-vectors and the global mean of all classes, respectively.

2.4. Multi-Layer Perceptron (MLP) Classifier

2.4.1. Architecture

Multi-Layer perceptron (MLP) represents a generalization of the single layer perceptron. A single layer perceptron forms half -plane decision regions, while multi – layer perceptron can form arbitrarily complex decision regions and can separatevarious input patterns [27]. Figure 5 shows a typical multi-layer perceptron neural network structure.
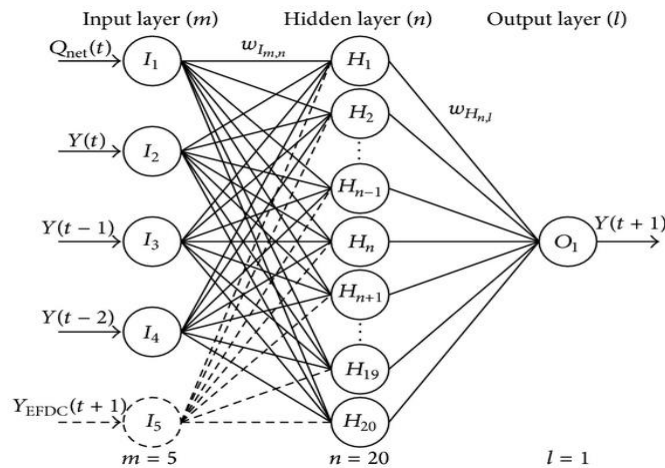
Figure 5: Multi-Layer Perceptron classifier.

In the context of biomedical signal processing, especially with application to EEG signals, the classification of data in feature spaces is often required. The objective of classification is to draw a boundary between two or more classes and to label them based on their measured feature, the task being to assign an input pattern represented by a feature vector to one or many pre-specified classes.

The MLP is a popular learningmachine for pattern classification[2,11, 28, 29]. In this paper, the goal is to classify different sets of EEG signals using MLP. MLP consists of the following layers.

Input layer: A layer of neurons that receives information from external sources and passes this information to the network for processing.

Hidden layer: A layer of neurons that receives information from the input layer and processes them in a hidden way. It has no direct connection withother layers within the system.
Output layer: A layer of neurons that receives processed information and sends output signal out of the system.

Bias: Acts on the neuron like an offset. The function of the bias is to provide a threshold for the activation of neurons. The bias input is connected to each of the hidden and output neurons in a network.

The ability to learn is a fundamental property of intelligence. Although a precise definition of learning is difficult to formulate, a learning process in the MLP can be viewed as the problem of updating connection weights so that a network can efficiently perform a specific task. The network usually must learn the connection weights from available training patterns. Performance is improved over time by iteratively updating the weights in the network. ANNs ability to automatically learn from examples makes them attractive and exciting. Instead of following a set of rules specified by human experts, ANNs appear to learn underlying rules from the given collection of representative examples. This is one of the major advantages of neural networks over traditional expert systems [30].

2.4.2. Learning process

There are three main learning paradigms: Supervised, Unsupervised, and hybrid. In Supervised learning, or learning with a "teacher", the network is provided with a correct answer (output) for every input pattern. Weights are determined to allow the network to provide answers as close as possible to the known correct answers. Examples of the supervised learning algorithms are the least mean square (LMS) algorithm and its generalization, known as the back-propagation algorithm[31-34]. In this work, the sigmoid function is used as activation function. The weights and bias are determined using the back-propagation algorithm, which is based on searching a minimum mean square error between the desired and actual outputs using gradient descent [29]. In contrast, unsupervised learning or learning without a teacher does not require a correct answer associated with each input pattern in the training data set. It explores the underlying structure in the data or correlation between patterns in the data and organizes patterns into categories from correlations. Kohonen network and K-means algorithm are examples of unsupervised learning.

Hybrid learning combines supervised and unsupervised learning. Part of the weights is usually determined through supervised learning, whileothers are obtained through unsupervised learning.

2.4.3. Back Propagation Algorithm

Backward of the output error

Learning by gradient descent
- Output layer:Correction guided by the error between the desired and the actual output;
- Hidden layers: Correction about the sensitivity (influence of the neuron on the error in the output layer).
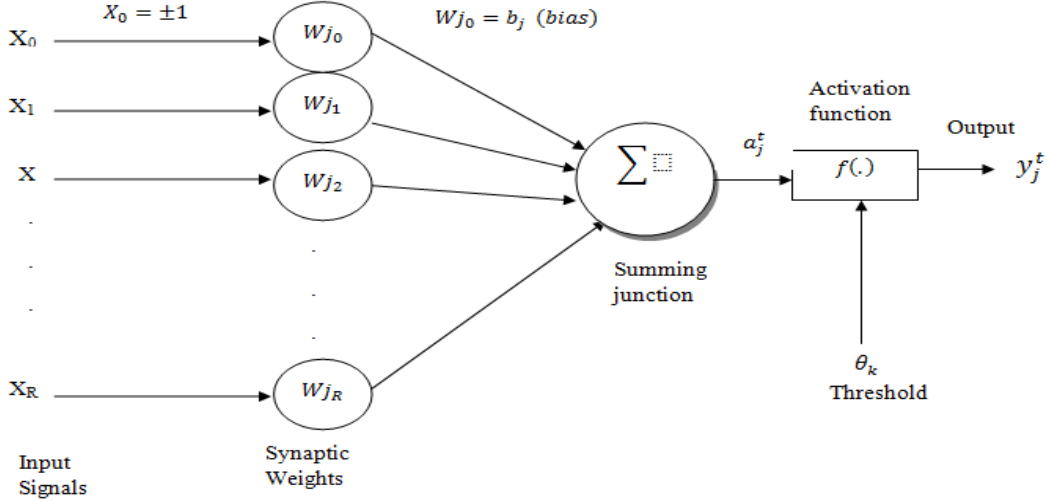

Figure 6:Mathematical Model of ANN.

Output values of the neurons

Value$y_j^t$of neuron j for the input dataX^t

$$y_j^t = f(a_j^t) = f\left(\sum_{i=1}^R \omega_{j,i} \cdot y_i^t + \omega_{j,0}\right) \tag{10}$$

f: activation function of the neuron;

$a_j^t = \sum_{i=1}^R \omega_{j,i} \cdot y_i^t + \omega_{j,0}$is the cumulative sum of input neurons;
$\omega_{j,i}$: weights of the link connecting neuron j to neuron i of the previous layer;
$\omega_{j,0}$: bias of neuron j;
y_i^t : output of neuron iof the previous layer for the input data X^t;
R: numberof neurons on the previous layer.

Output error to the output layer

A set of data $\chi = \{x^t, r^t\}_{t=1}^N$with $r^t = [r_1^t r_2^t \dots r_K^t]^T = \begin{bmatrix} r_1^t \\ r_2^t \\ . \\ . \\ r_K^t \end{bmatrix}$

Where$r_j^t = 1 \ if x^t \epsilon C_j, otherwiser_j^t = 0$

Error observed for the data x^ton neuron j on the output layer

$$e_j^t = r_j^t - y_j^t \tag{11}$$

Quadratic error observed for the data x^t on all the K neurons of the output layer (one neuron per class).

$$E^t = \frac{1}{2}\sum_{j=1}^{K}\left(e_j^t\right)^2 \tag{12}$$

Quadratic mean error observed for the data set χ.

$$E = \frac{1}{N}\sum_{t=1}^{N}E^t \tag{13}$$

Correction of the error for the output layer

Correction of weights through descent gradient of the quadratic mean error

$$\Delta\omega_{j,i} = -\eta\frac{\partial E}{\partial\omega_{j,i}} = -\eta\frac{\partial\left(\frac{1}{N}\sum_{t=1}^{N}E^t\right)}{\partial\omega_{j,i}} \tag{14}$$

$$\Delta\omega_{j,i} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial\omega_{j,i}}$$

The error of neuron j depends on neurons of the previous layer
Expanding using the chain derivative rule. $\left(\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y}\frac{\partial y}{\partial x}\right)$

$$\frac{\partial E^t}{\partial\omega_{i,j}} = \frac{\partial E^t}{\partial e_j^t}\frac{\partial e^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j}\frac{\partial a_j^t}{\partial\omega_{j,i}} \tag{15}$$

$$\frac{\partial E^t}{\partial\omega_{j,0}} = \frac{\partial E^t}{\partial e_j^t}\frac{\partial e^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j}\frac{\partial a_j^t}{\partial\omega_{j,0}} \tag{16}$$

Calculation of the partialderivatives

Expanding with the sigmoid activation function $y_j^t = \frac{1}{1+\exp\left(-a_j^t\right)}$ (17)

$$\frac{\partial E^t}{\partial e_j^t} = \frac{\partial}{\partial e_j^t}\left[\frac{1}{2}\sum_{j=1}^{K}\left(e_j^t\right)^2\right] = e_j^t \tag{18}$$

$$\frac{\partial e_j^t}{\partial y_j^t} = \frac{\partial\left[r_j^t-y_j^t\right]}{\partial y_j^t} = -1 \tag{19}$$

$$\frac{\partial y_j^t}{\partial a_j^t} = \frac{\partial\left[\frac{1}{1+exp\left(-a_j^t\right)}\right]}{\partial a_j^t} = \frac{exp\left(-a_j^t\right)}{\left[1+exp\left(-a_j^t\right)\right]^2} = \left[\frac{1}{1+exp\left(-a_j^t\right)}\right]\left[1 - \left\{\frac{1}{1+exp\left(-a_j^t\right)}\right\}\right] \tag{20}$$

$$\frac{\partial y_j^t}{\partial a_j^t} = \frac{1}{1+exp\left(-a_j^t\right)} * \frac{e^{-a_j^t}+1-1}{1+exp\left(-a_j^t\right)} = y_j^t\left(1-y_j^t\right) \tag{20}$$

$$\frac{\partial y_j^t}{\partial a_j^t} = y_j^t\left(1-y_j^t\right) \tag{20}$$

$$\frac{\partial a_j^t}{\partial\omega_{j,i}} = \frac{\partial}{\partial\omega_{j,i}}\sum_{i=1}^{R}\omega_{j,i}y_i^t + \omega_{j.0} = y_i^t \tag{21}$$

$$\frac{\partial a_j^t}{\partial\omega_{j,0}} = 1 \tag{22}$$

Learning for the output layer

Adjusting the weights of the output layer

$$\Delta\omega_{j,i} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial\omega_{j,i}} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial e_j^t}\frac{\partial e_j^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j^t}\frac{\partial a_j^t}{\partial\omega_{j,i}} \tag{23}$$

$$\Delta\omega_{j,i} = +\frac{\eta}{N}\sum_{t=1}^{N} e_j^t y_j^t(1-y_j^t)y_i^t \tag{24}$$

$$\Delta\omega_{j,0} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial\omega_{j,0}} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial e_j^t}\frac{\partial e_j^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j^t}\frac{\partial a_j^t}{\partial\omega_{j,0}} \tag{25}$$

$$\Delta\omega_{j,0} = \frac{\eta}{N}\sum_{t=1}^{N} e_j^t y_j^t(1-y_j^t) \tag{26}$$

Delta Rule

Let a delta δ_j^t, which corresponds to the local gradient of neuron j for the input data X^t.

$$\delta_j^t = e_j^t y_j^t(1-y_j^t) \tag{27}$$

$$\text{Thus}\Delta\omega_{j,i} = \frac{\eta}{N}\sum_{t=1}^{N}\delta_j^t y_i^t \tag{28}$$

$$\text{And }\Delta\omega_{j,0} = \frac{\eta}{N}\sum_{t=1}^{N}\delta_j^t \tag{29}$$

Correction of the error for the hidden layers

Gradient of the error for the hidden layers.

$$\frac{\partial E^t}{\partial\omega_{j,i}} = \frac{\partial E^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j^t}\frac{\partial a_j^t}{\partial\omega_{j,i}} \tag{30}$$

Only $\frac{\partial E^t}{\partial y_j^t}$ changes, while $\frac{\partial y_j^t}{\partial a_j^t}$ *and* $\frac{\partial a_j^t}{\partial\omega_{j,i}}$ are the same as on the output layer.

The error for a neuron of the hidden layer depends on the error of the K neurons of the next layer. (Back-propagation of the error)

$$E^t = \frac{1}{2}\sum_{k=1}^{N}(e_k^t)^2 \tag{31}$$

$$\frac{\partial E^t}{\partial y_j^t} = \frac{\partial}{\partial y_j^t}\left[\frac{1}{2}\sum_k(e_k^t)^2\right] = \sum_k e_k^t\frac{\partial e_k^t}{\partial y_j^t} \tag{32}$$

$$\frac{\partial E^t}{\partial y_j^t} = \sum_k e_k^t\frac{\partial e_k^t}{\partial a_k^t}\frac{\partial a_k^t}{\partial y_j^t} = \sum_k e_k^t\frac{\partial[r_k^t-y_k^t]}{\partial a_k^t}\frac{\partial[\sum_l\omega_{k,l}y_k^t+\omega_{k,0}]}{\partial y_j^t} \tag{33}$$

$$\frac{\partial E^t}{\partial y_j^t} = \sum_k e_k^t[-y_k^t(1-y_k^t)]\omega_{k,j} \tag{33}$$

$$\delta_k^t = e_k^t[y_k^t(1-y_k^t)] \tag{34}$$

$$\frac{\partial E^t}{\partial y_j^t} = -\sum_k\delta_k^t\omega_{k,j} \tag{35}$$

Correction of the corresponding error

$$\frac{\partial E^t}{\partial\omega_{j,i}} = \frac{\partial E^t}{\partial y_j^t}\frac{\partial y_j^t}{\partial a_j^t}\frac{\partial a_j^t}{\partial\omega_{j,i}} = -\left[\sum_k\delta_k^t\omega_{k,j}\right]y_j^t(1-y_j^t)y_i^t \tag{36}$$

$$\delta_j^t = y_j^t(1-y_j^t)\sum_k\delta_k^t\omega_{k,j} \tag{37}$$

$$\Delta\omega_{j,i} = -\eta\frac{\partial E}{\partial\omega_{j,i}} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial\omega_{j,i}} = \frac{\eta}{N}\sum_{t=1}^{N}\delta_j^t y_i^t \tag{38}$$

$$\Delta\omega_{j,0} = -\eta\frac{\partial E}{\partial\omega_{j,0}} = -\frac{\eta}{N}\sum_{t=1}^{N}\frac{\partial E^t}{\partial\omega_{j,0}} = \frac{\eta}{N}\sum_{t=1}^{N}\delta_j^t \tag{39}$$
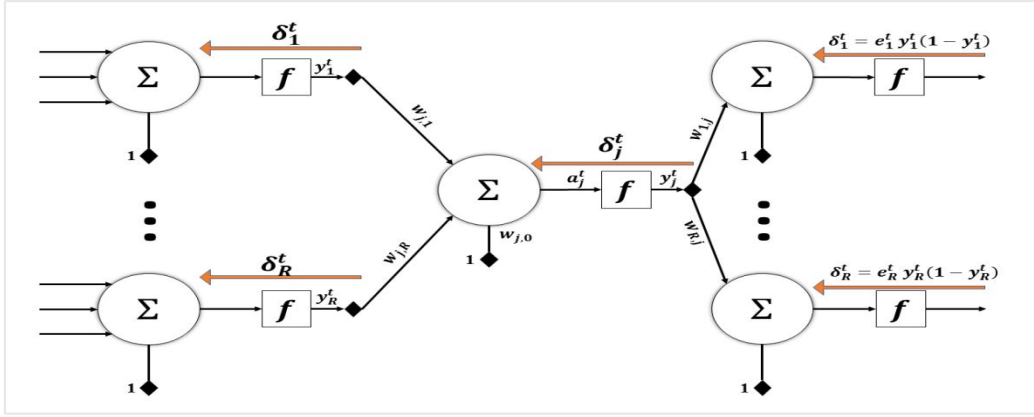
Figure 7: Back-propagation of error.

The back-propagation algorithm can be summarized by the following steps:

Step1: Normalize the training data $X_i^t \varepsilon [-1,1]$ and the desired outputs $\tilde{r}_i^t \varepsilon \{0,05 ; 0,95\}$;

Step 2: Initialize the weights and bias to small random values, $\omega_{j,i} \varepsilon \mathcal{U}(-0,5 , 0,5)$;

Step3: Select the structure of the network (such as the number of hidden layers and number of neurons for each layer);

Step4: Choose activation functions for the neurons. These activation functions can be uniform, or they can be different for different layers;

Step5: Select the training pair from the training set. Apply the input vector to the network input and propagate the signal forward through the network;

Step6: Calculate the output of the network based on the initial weights and input set;

Step7: Calculate the error between network output and desired output;

$$e_j^t = \tilde{r}_j^t - y_j^t ; \quad j = 1 \ldots \ldots , K ; t = 1, \ldots \ldots \ldots , N \tag{40}$$

Step8: Propagate error backward and adjust the weights in such a way that minimizes the error. Start from the output layer and go backward to the input layer;

$$\omega_{j,i} = \omega_{j,i} + \Delta\omega_{j,i} = \omega_{j,i} + \frac{\eta}{N} \sum_t \delta_j^t y_i^t \tag{41}$$

$$\omega_{j,0} = \omega_{j,0} + \Delta\omega_{j,0} = \omega_{j,0} + \frac{\eta}{N} \sum_t \delta_j^t \tag{42}$$

Where the local gradient or Delta is defined as:

$$\delta_j^t = \begin{cases} e_j^t y_j^t \left(1 - y_j^t\right), & \text{if } j \text{ belongs to the output layer} \\ y_j^t \left(1 - y_j^t\right) \sum_k \delta_k^t \omega_{k,j}, & \text{if } j \text{ belongs to the hidden layer} \end{cases} \tag{43}$$

Step9: Repeat steps 5-8 for each vector in the training set until the error for the set is lower than the required minimum error.

After enough repetitions of these steps the error between the actual output and the target output should be reduced to an acceptable value, and the network is said to be trained. At this point the network can be used in the recall or generalization phase where the weights are not changed.

2.5. Classification Problems and Performances Evaluation

2.5.1. Terminology and Standard Measures

The American Heart Association (AHA) recommends the use of four measures to evaluate procedures for diagnosing CAD [35]. Since these measures are useful in other areas of diagnosis as well, we will be using them in evaluating epilepsy diagnosis system.

$$\text{Sensitivity} \ldots . TPF = \frac{TP*100}{TP+FN} \tag{44}$$

$$\text{Specificity} \ldots. TNF = \frac{TN*100}{TN+FP} \tag{45}$$

$$PA = Sensitivity * P(D) + Specificity * [1 - P(D)] \tag{46}$$

$$PV = \frac{Sensitivity*P(D)}{Sensitivity*P(D)+(100-Specificity)*[1-P(D)]} \tag{47}$$

$$ACC = \frac{TP+TN}{TP+FN+TN+FP} * 100 \tag{48}$$

Where TP stands for true positive, FN stands for false negative, TN stands for true negative, and FP stands for false positive.

Sensitivity, or true-positive fraction (TPF), is the probability of a patient who is suffering from an epilepsy to be diagnosed as such.

Specificity, or true negative fraction (TNF), is the probability that a healthy individual is diagnosed as such by an epilepsy diagnosis mechanism.

PA is the predictive accuracy or overall percentage of correct diagnosis.

PV is the predictive value of a positive test or the percentage of those who have the disease and have tested positive for it.

P(D) is the a priori probability of a patient who is referred to the diagnosis procedure having epilepsy.

Accuracy (ACC): sum of correctly classified patterns ($TP + TN$)/sum of correctly and non-correctly classified patterns ($TP + FN + TN + FP$).

In addition to TPF and TNF, we define two other related values.

False positive fraction (FPF) is the probability of a healthy patient being incorrectly diagnosed as suffering from epilepsy.

False negative fraction (FNF) is the probability that a patient who is suffering from epilepsy will be incorrectly diagnosed as healthy. In this way, the following relations can be established:

$$FPF = 1 - TNF \tag{49}$$

$$FNF = 1 - TPF \tag{50}$$

In this paper, five classification problems that are closely related to the clinical applications are designed from the entire database.

Table 3. Classes and corresponding number of EEG for five classification problems.

Classification problem	Classes	Number of EEG segments
A – E	Normal (A)	100
	Ictal (E)	100
B – E	Normal (B)	100
	Ictal (E)	100
AB - E	Normal (A, B)	200
	Ictal (E)	100
AB - CD	Normal (A, B)	200
	Pre-ictal (C, D)	200
CD-E	Pre-ictal (C, D)	200
	Ictal (E)	100

The sensitivity, specificity, and accuracy parameters are used as performance metrics of the MLP for each classification problem [4].

3. RESULTS AND DISCUSSION

This section presents and discusses the experimental results of the proposed model in order to determine its capability and consistency. All experiments have been carried out in JAVA. In this study, all the database is used to perform different classification problems. Firstly, EEG rhythms are separated using WT and the frequency distribution of the EEG signal is analyzed in terms of the areas under the frequency spectrum of each wavelet coefficient. According to their supervision, it is observed that areas analysis of the database is representative for the various classes of signals but still presents redundancy. Therefore, the LDA is applied to select and extract suitable, easily viewed and interpreted two-dimension features. Figure 5 presents how these extracted 2-dimension feature vectors represent the amount of changes in frequency distribution and carry discriminative and representative information about their classes for different classification problems. Then, these feature vectors can be used as benchmark of classification research through the proposed MLPN classifier. Compared to others, these feature vectorsare extremely discriminative since more of the previous proposed models use complex kernel machines after feature extraction. For different classification problems, the first 40% of each SET are used to determine the useful linear discriminants for the feature extraction process, the remaining 60 % of each set is used for the classification process.
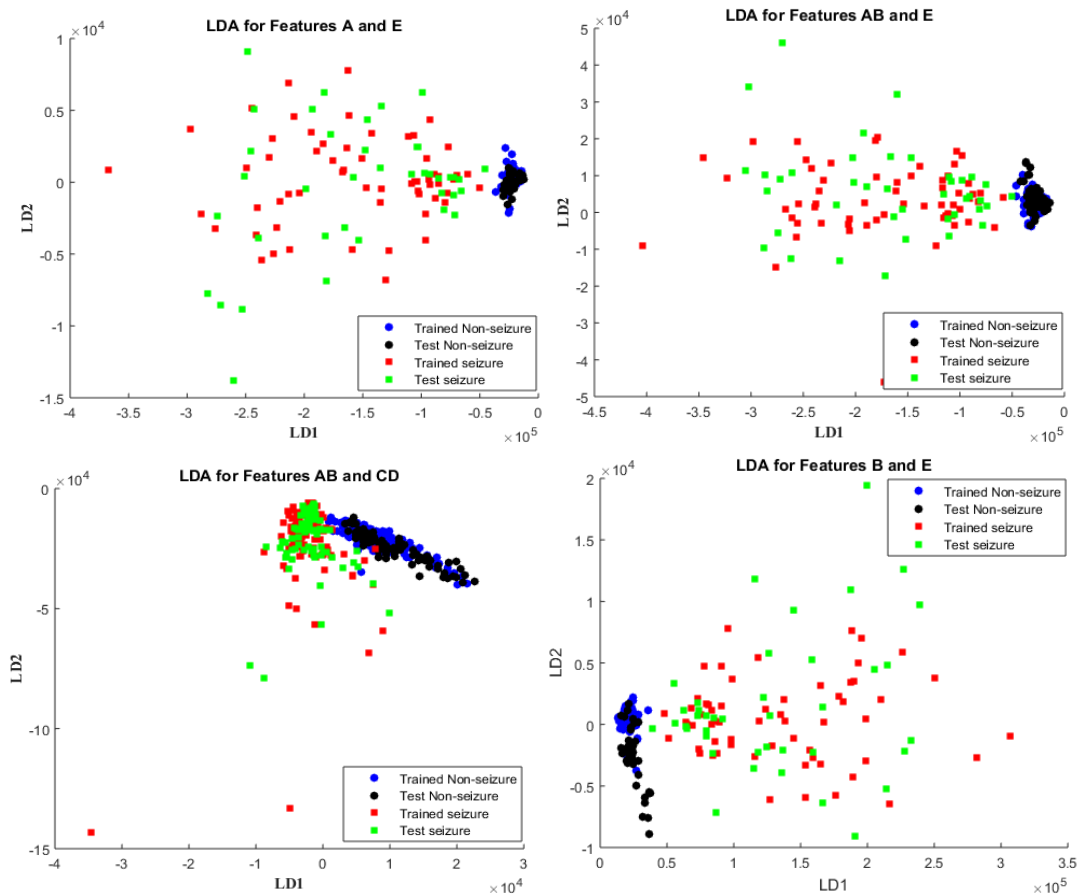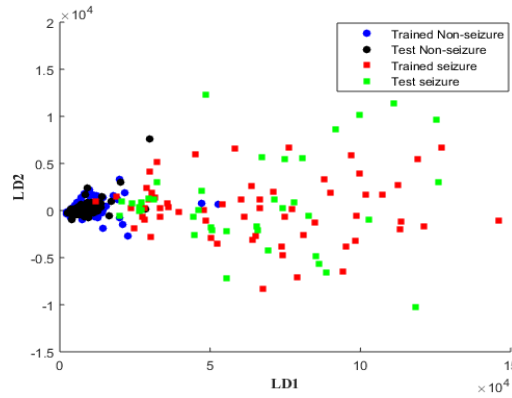
Figure 5: Scatter plot of LDA features for different classification problems.

Therefore, the performances of the MLPNN for the five classification problems are evaluated and summarized in Table 4 in terms of sensitivity, specificity and accuracy. In the first two classifications $\eta = 0.8$, and $\eta = 0.6$ for the rest. The first 70% of each SET are used as the training set and the last 30% is divided into two parts; 15 % as the validation set and 15 % for the testing set. In order to achieve a good generalization or convergence, the input features and the output target are normalized.
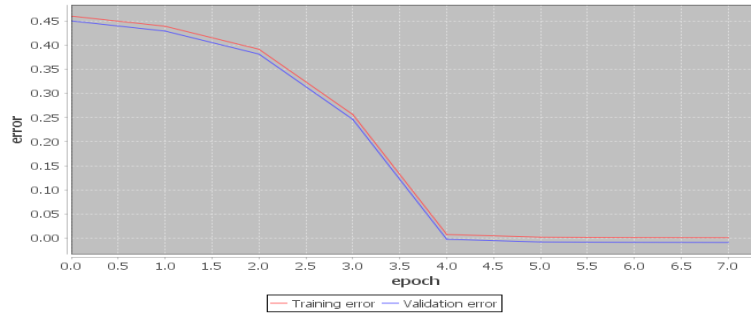

A-E (η= 0.8)

B-E (η= 0.8)

AB-CD (η= 0.8)
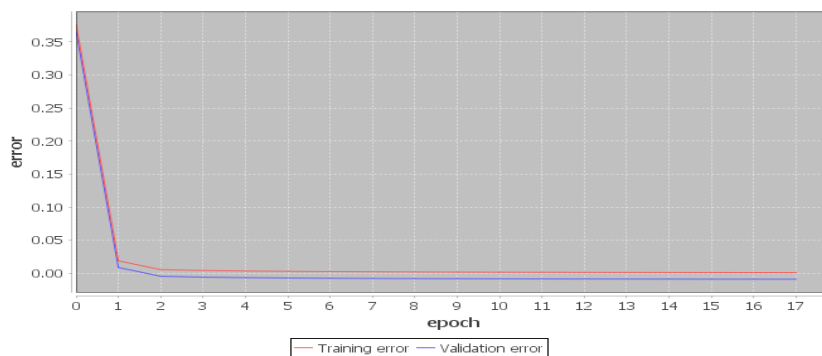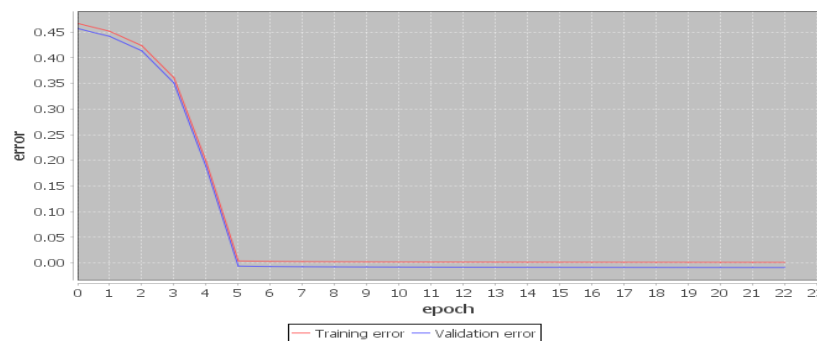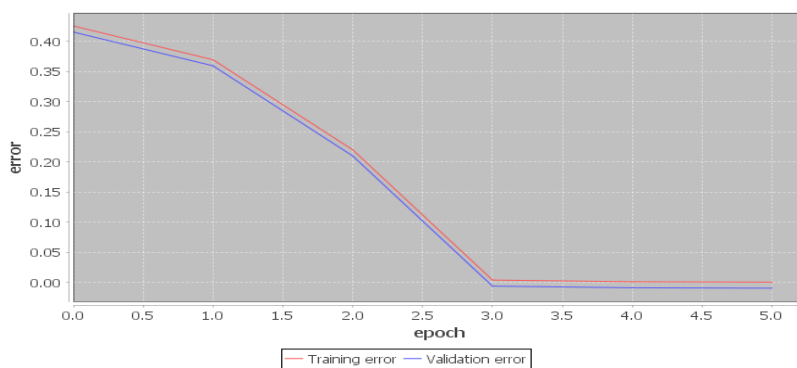
CD-E (η= 0.6)

Figure 6: Convergence of training and validation error of the MLPNN with 1, 12, 12 and 2 neurons at the input layer, first hidden layer, second hidden layer and output layer, respectively.

Table 4: Performances of the MLPNN for five classification problems.

Classification Problem	MLPNN Performances (%)		
	Sensitivity	Specificity	Accuracy
A - E	100	100	100
B - E	100	100	100
AB - E	100	100	100
CD - E	100	100	100
AB - CD	100	100	100

For A-E, B-E, AB-E, CD-E and AB-CD classification problems, the proposed model achieved a maximum accuracy of 100%. Then, for this clinical diagnosis, the proposed model can be used, and good decision can be taken out.

Compared to other works, in general, for different experiments; the model proposed in this paper achieves a high accuracy. The MLPNN architecture used in this work has a lower computational cost and is faster to compute. Then, this WT-LDA-MLPNN model can be used for EEG signals classification problems since highly discriminative features are extracted and high accuracy is obtained.

4. CONCLUSION

In this paper, WT has been applied for time-frequency decomposition and their rhythm distributions have been analyzed using areas under the frequency spectrum of each wavelet coefficient. Furthermore, LDA has been used to extract discriminative 2-dimension feature vectors as input of a MLPNN. In terms of their sensitivity, specificity and accuracy, the performances for five classification problems were derived. The result of EEG signal classification through MLPNN has demonstrated that LDA feature extraction can improve the performance of classifiers with respect to reducing the computational cost. According to this result, the proposed model can serve as a promising alternative for intelligent diagnosis system.

REFERENCES

[1]   SaeidSanei and J.A. Chambers., [2007] "EEG signal Processing". John Wiley & Sons.
[2]   Blume W., Lüders H., Mizrahi E., Tassinari C., van Emde Boas W., Engel J. [2001], *"Glossary of Descriptive Terminology for Ictal Semiology: Report of the ILAE Task Force on Classification and Terminology"*, *Epilepsia,* Vol. 42, No. 9, pp. 1212-1218
[3]   Subasi Abdulhamit [2007], ''*EEG Signal Classification using Wavelet Feature Extraction and a Mixture of Expert Model''*, Elsevier Transactions on Expert System with Applications, Vol. 32, No. 5, pp. 1084-1093.

[4]   K. Fu, J.Qu, Y. Chai, and Y. Dong [2014] "Classification of seizure based on the time – frequency image of EEG signals using HHT and SVM". Biomedical Signal Processing and Control vol -13, pp: 15-22.

[5]   Adeli H., Zhou Z. and Dadmehr N. [2003] *"Analysis of EEG Records in an Epileptic Patient Using Wavelet Transform",* Journal of Neuroscience Methods, Vol. 123, pp. 69-87.

[6]   Subasi A., Gursoy M. I. [2010], *''EEG Signal Classification Using PCA, ICA, LDA and Support Vector Machines''*, Expert Systems with Applications, Vol. 37, pp. 8659-8666.

[7]   Tzallas A. T., Tsipouras M. G., Fotiadis D. I. [2007], *"Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks",*Comput. Int. Neurosci., pp. 1-13. DOI: 10.1155/2007/80510.

[8]   Pachori R. B., Bajaj V. [2008], *''Discrimination between Ictal and Seizure-Free EEG Signals Using Empirical Mode Decomposition''*, Res. Lett. Signal Process. DOI: 10.1155/2008/293056

[9]   Sivasankari N. and Thanushkodi K. [2009],*"Automated Epileptic Seizure Detection in EEG Signals Using FastICA and Neural Network"*, Int. J. Advance. Soft Comput. Appl., Vol. 1, No. 2, November. ISSN: 2074-8523

[10]   Ocak H. [2009], *''Automatic Detection of Epileptic Seizures in EEG Using Discrete Wavelet Transform and Approximate Entropy''*, Expert Systems with Applications, Vol. 36, No. 2, pp. 2027-2036.

[11]   Ling Guo, Daniel Rivero, Julián Dorado, Juan R. Rabunal, Alejandro Pazos [2010], *"Automatic Epileptic Seizure Detection in EEG Based on Line Length Feature and Artificial Neural Networks",* Journal of Neuroscience Methods, Vol. 191, pp. 101-109. DOI: 10.1016/j.jneumeth.2010.05.020.

[12]   Pachori R. B. and Bajaj V. [2011], *''Analysis of Normal and Epileptic Seizure EEG Signals Using Empirical Mode Decomposition''*, Comput. Methods Progr. Biomed., Vol. 104, pp. 373-381.

[13]   OrhanUmut, HekimMahmut, Ozer Mahmut [2011], *"EEG Signals Classification Using the K-Means Clustering and a Multilayer Perceptron Neural Network Model"*, Expert Systems with Applications, Vol. 38, pp. 13475-13481..

[14]   Nicolaou N., Georgiou J. [2012], *''Detection of Epileptic Electroencephalogram Based on Permutation Entropy and Support Vector Machines''*, Expert Systems with Applications, Vol. 39, No. 1, pp. 202-209.

[15]   Saching Garg, Rakesh Narvey [2013], *''Denoising and Feature Extraction of EEG Signal Using Wavelet Transform''*, International Journal of Engineering Science and Technology (IJEST), Vol. 5, No. 6, pp. 1249-1253, June. ISSN: 0975-5462.

[16]   Rosenblatt Mariel et al. [2014], *''A Quantitative Analysis of an EEG Epileptic Record Based on Multiresolution Wavelet Coefficients''*, Entropy, Vol. 16, pp. 5976-6005, 2014.

[17]   Kai Fu, Qu J., Chai Y., Dong Y. [2014], *''Classification of Seizure Based on the Time-Frequency Image of EEG Signals Using HHT and SVM''*, Biomedical Signal Processing and Control, Vol. 13, pp. 15-22.

[18]   Ashwini Nakate, P. D. Bahirgonde [2015], ''*Feature Extraction of EEG Signal UsingWavelet Transform*'', International Journal of Computer Applications,Vol. 124, No. 2, August.

[19]   Pushpendra Singh, S. D. Joshi,·R. K. Patney, Kaushik Saha [2016], *"Fourier-Based Feature Extraction for Classification of EEG Signals Using EEG Rhythms",*Circuits Syst Signal Process, Vol. 35, pp. 3700-3715. DOI: 10.1007/s00034-015-0225-z.

[20]   Ramakrishnan J, Bommanna Raja Kanagaraj [2017], *"Analysis of Non-seizure and Seizure Activity Using Intracranial EEG Signals and Empirical Mode Decomposition Based Approximate Entropy",*Biomedical Research, Vol. Special Issue. ISSN: 0970-938X.

[21]   Sandeep Kumar Satapathy, SatchidanandaDehuri, Alok Kumar Jagadev [2017], *"EEG Signal Classification Using PSO Trained RBF Neural Network for Epilepsy Identification"*, Informatics in Medicine Unlocked, Vol. 6, pp.1-11.

[22]   Mohd Zuhair, Sonia Thomas [2017], *''Classification of Patient by Analyzing EEG Signal Using DWT and Least Square Support Vector Machine''*, Advances in Science, Technology and Engineering Systems Journal, Vol. 2, No. 3, pp. 1280-1289, 2017. ISSN: 2415-669.

[23]   Subasi Abdulhamit, Kevric Jasmin, Canbaz M. Abdullah [2017], *"Epileptic Seizure Detection Using Hybrid Machine Learning Methods"*, Neural Comput and Applic. DOI: 10.1007/s00521-017-3003-y.

[24]   Wang Lina, XueWeining, Li Yang, Luo Meilin, Huang Jie, Cui Weigang and Huang Chao [2017], *"Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis"*, Entropy, Vol. 19, No. 222. DOI: 10.3390/e19060222

[25]   Jérôme Landre et al., [2005] "Analyse multiresolution pour la recherché et l'indexation d'image par le contenu dans les bases de donnees image". Application a la base d'images paléontologiques Trans' Tyfipal; Interface homme-machine. Thèse de Doctorat de l'Université de Bourgogne.

[26] PulomaPramanick et al., [2013] 'Classification of Electroencephalogram (EEG) signal based on Fourier Transform and neural network'. Mémoire de recherche au Génie Electrique de l'Institut National de Technologie de Rourkela.

[27] Nicolaou N., Georgiou J. [2012], *''Detection of Epileptic Electroencephalogram Based on Permutation Entropy and Support Vector Machines''*, Expert Systems with Applications, Vol. 39, No. 1, pp. 202-209.

[28] Andrzejak R. G., Lehnertz K., Mormann F., Rieke C., David P. and Elger C. E. [2001], *''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, No. 6. DOI: 061907.

[29] Daubechies I. [1992], ''*Ten Lectures on Wavelets*'', Industrial and Applied Mathematics, Philadelphia

[30] Mallat S. [1989], ''*A Theory for Multi-resolution Signal decomposition: The wavelet Representation*'', IEEE Transaction on Analy. Mach. Intell., Vol. 11, No. 7, pp. 674-693.

[31] M. Zecca, S. Micera, MC. Carrozza, and P. Dario [2002] "Control of multifunctional prosthetic hands by processing the electromyography signal'. Critical Reviews in Biomedical Engineering, citeeer, 30(4-60; 459.

[32] A. Bashashati, M. Fatourechi, R. K Ward, and G. E Birch [2007]. "A Survey of signal processing algorithms in brain- computer interface based on electrical brain signals". Journal of Neural engineering 4; R32.

[33] Ali Zilouchian and Mohammad Jamshidi : Intelligent control systems using soft computing methodologies.p. cm.Includes bibliographical references and index.ISBN 0-8493-1875-01. Intelligent control systems—Data processing. 2. Soft computing. TJ217.5 .I5435 2001 629.89′ 0285′63—dc21

[34] Miller A. S., Blott B. H. and Hames T. K. [1992], *"Review of Neural Network Applications in Medical Imaging and Signal Processing",*Medical and Biological Engineering and ComputingVol. 30, pp. 449-464.

[35] Fausett L. [1994], *"Fundamentals of Neural Networks Architectures, Algorithms, and Applications"*, Prentice Hall, Englewood Cliffs, NJ.

[36] Anil K. Iain; Jianchang Mao, et K.M. Mobraddin: Artificial Neural Networks: A Tutorial 1996

[37] Widrow. B and Lehr, M. A., " Thirty Years of Adaptive Neural Networks: Perceptron, MADALINE , and Back propagation. Proc. of the IEEE, vol. 78, 1415-1442, 1990

[38] Rumelhart, D.E., Hinton G.E., and Williarns, R.J., Learning Internal Representations by Error Propagation, Parallel Data Processing , vol. 1, chap 8 , The MIT Press, Cambridge, MA , 1986

[39] Specht, D.F., A General Regression Neural Network*, IEEE Trans. Of Neural Networks,* Vol. 2, 568□576, 1991

[40] Wasserman P.D., *Advanced Methods in Neural Computing*,VanNostrand Reinhold, New York, 1993

[41] Gibbons, R.J., et al., ACC/AHA Guidelines for Exercise Testing, A Report of American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Exercise Testing), J. of the Am. Coll., of Cardiol., 30(1) 260-311, July 1997

AUTHORS

Atangana Romain was born in 1967 in Ngoumou-Cameroon. In 1997, he graduated from the Higher technical Teacher's Training College (ENSET)-University of Douala with DIPET 1 (bachelor's degree in electrical and Electronics Engineering). In 2002 he obtained the DIPET 2 in the same institution. He obtained a DEA (Master of Sciences with Thesis) at "Ecole DoctoralePhysoique et Sciences de l'Ingénieur" in Electronics, Electrotechnic, Automatic and Teledetection (EEAT) in 2007 at the University of Douala-Cameroon. He is currently the Head of service of distance learning at the Higher Teacher's Training College (ENS) of Bertoua-University of Ngaoundéré. In parallel to his job, Mr. Atangana Romain is following studies and research for a PhD thesis. His research interests are telemedicine, signal and image processing, artificial intelligence and automation, medical physics, and biomedical engineering. He is a member of "Unité de Recherche d'Automatique et d'InformatiqueAppliquée (UR-AIA)" of the IUT-FV of Bandjoun of the University of Dschang-Cameroon, and "Unité de Recherche de Matière Condensée et de Traitement de Signal (UR-MACETS)" of the Faculty of Science of the University of Dschang-Cameroon.

Tchiotsop Daniel was born in 1965 in Tombel - Cameroon. He graduated in Electromechanicalengineering from "Ecole Nationale Supérieure Polytechnique (ENSP)" of Yaoundé-Cameroon in 1990, he obtained a MS degree in Solid Physics in 1992 from the Faculty of Science of the University of Yaoundé I, a MS degree in Electrical Engineering and Telecommunication in 2003 from ENSP-Yaoundé and a PhD at INPL ("Institut National Polytechnique de Lorraine"), Nancy-France, in 2007. Dr TCHIOTSOP teaches in the Department of Electrical Engineering of "FOTSO Victor University Institute of Technology" – University of Dschang since 1999 where he is actually the Head of Department. He is with "Laboratoired'Automatique et d'InformatiqueAppliquée (LAIA)" where his main items of research include Biomedical Engineering, Biomedical signal and image processing, Telemedicine and intelligent systems.Dr TCHIOTSOP is partner with "Centre de Recherche enAutomatique de Nancy (CRAN)-Université de Lorraine", France, "Laboratoired'Electronique et du Traitement de Signal (LETS)-ENSP, University of Yaoundé 1", and "Unité de Recherche d'Automatique et d'InformatiqueAppliquée (UR-AIA)" of the IUT-FV of Bandjoun of the University of Dschang-Cameroon.

Kenne Godpromesse was born in Mbouda-Cameroon in 1967. He obtaineda B.S. degree in electromechanical engineering and a M.S. degree in mechanical engineering from "Ecole Nationale Supérieure Polytechnique, Université de Taoundé I", Yaoundé-Cameroon in 1991 and 1994, respectively, and a PhD degree in control theory from "Université Paris XI", Orsay, France, in 2003. Since 1996, he has been with the UniversityofDschang, Daschang-Cameroon, where he is currently a Senior Lecturer. From 1996 to 2005, he has been the Head of the Department of Electrical Engineering, Fotso Victor University Institute of Technology,Bandjoun-Cameroon. He was a Guest Researcher at "Laboratoire des Signaux et Systèmes (L2S)", "Centre National de la Recherche Scientifique", "L'Ecoled'Electricité (CNRS-SUPELEC)", "Université Paris XI", and the Energy Department, SUPELEC, from September 2005 to January 2006 and from June to July 2006. Since April 2006, he has been a founder of "Unité de Recherche d'Automatique et d'InformatiqueAppliquée (UR-AIA)" of the IUT-FV of Bandjoun of the University of Dschang-Cameroon. His current research interests include identification and control of nonlinear systems using variable structure and neural network with applications in electromechanical systems.

DjoufackNkengfack Laurent Chanel was born in 1991 in Nanga-Eboko-Cameroon. In 2014, he obtained abachelor's degree in physics option Electronics, Electrotechnic and Automatic at the University of Dschang-Cameroun. In 2016 he was graduating from the same institution with aMaster of Sciences with thesis in Physics option Electronics, Electrotechnic and Automatic. In 2018 he obtained a DIPET 2 in Electrical Engineering option Electronics at the Higher Teacher's Training College of Bambili-University of Bamenda-Cameroon. He is currently a Technical High School teacher in electronics engineering. In parallel to his job, Mr. Djoufack is following studies and research for a PhD thesis. His research interests are telemedicine, signal and image processing, artificial intelligence, medical physics, and biomedical engineering. He is a member of "Unité de Recherche d'Automatique et d'InformatiqueAppliquée (UR-AIA)" of the IUT-FV of Bandjoun of the University of Dschang-Cameroon, and "Unité de Recherche de Matière Condensée et de Traitement de Signal (UR-MACETS)" of the Faculty of Science of the University of Dschang-Cameroon.