UNIVERSITY OF TURKISH AERONAUTICAL ASSOCIATION
INSTITUTE OF NATURAL AND APPLIED SCIENCES

AUTOMATIC FEATURE EXTRACTION OF EEG SIGNALS USING NEURAL NETWORKS AND TIME-FREQUENCY ANALYSIS

MASTER THESIS

Omar Saadi Fathi ALSHEAR

Electrical and Electronics Engineering Department

Master Thesis Program

DECEMBER 2017

UNIVERSITY OF TURKISH AERONAUTICAL ASSOCIATION
INSTITUTE OF NATURAL AND APPLIED SCIENCES

AUTOMATIC FEATURE EXTRACTION OF EEG SIGNALS USING NEURAL NETWORKS AND TIME-FREQUENCY ANALYSIS

MASTER THESIS

Omar Saadi Fathi ALSHEAR

1406030041

Electrical and Electronics Engineering Department

Master Thesis Program

Supervisor: Asst. Prof. Dr. Hassan SHARABATY

Omar Saadi Fathi ALSHEAR, having student number 1406030041 and enrolled in the Master Program at the Institute of Science and Technology at the University of Turkish Aeronautical Association, after meeting all of the required conditions contained in the related regulations, has successfully accomplished, in front of the jury, the presentation of the thesis prepared with the title of: "Automatic Feature Extraction of EEG Signals Using Neural Networks and Time-Frequency Analysis."

Supervisor : **Asst. Prof. Dr. Hassan SHARABATY**
University of Turkish Aeronautical Association ....

Jury Members : **Asst. Prof. Dr. Abdül Kadir GÖRÜR**
Çankaya University ....

: **Asst. Prof. Dr. Javad RAHEBİ**
University of Turkish Aeronautical Association ....

: **Asst. Prof. Dr. Hassan SHARABATY**
University of Turkish Aeronautical Association ....

Thesis Defense Date: 05.12.2017

STATEMENT OF NON-PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

15.11.2017

Omar Saadi ALSHEAR

DEDICATION

I dedicate this work to,

To the sake of Allah, my Creator and my Master.

To my great teacher and messenger, Mohammed (May Allah bless and grant him), who taught us the purpose of life.

To my parents who gave me the strength and patience enough to finish this work.

ACKNOWLEDGEMENTS

This is a nice opportunity to show deep thankfulness to all those who helped and supported me in completing this piece of work. First of all, I would like to show sincere thankful to my supervisor Asst. Assoc. Dr. Hassan Sharabaty, who willingly shared his knowledge, experience, and ideas with me, where without his excellent courses, guidance, and remarkable patience I don't think I would ever have finished this work. I would like to thank all of the academic staff in the Electrical and Electronic Engineering Department for allocating their valuable time and effort during these years of study. I want to express my sincere gratitude to my, wife, brothers, children, and all my family for their complete reliance, unconditional love, and moral support throughout my academic years.

December 2017 Omar Saadi ALSHEAR

TABLE OF CONTENT

LIST OF ABBREVIATION

DWT	**:**	Discrete Wavelet Transform
EEG	**:**	Electroencephalogram
ANN	**:**	Artificial Neural Network
TDF	**:**	Time Domain Feature
FDF	**:**	Frequency Domain Feature
SVM	**:**	Support Vector Machine
MLPNN	**:**	Multilayer Perceptrun Neural Network
BP	**:**	Back Propagation
LM	**:**	Levenberg-Marquardt
DWT	**:**	Discrete Wavelet Transform
ME	**:**	Mixture of Experts
LS-SVM	**:**	Least Square Support Vector Machine
DCT	**:**	Discrete Cosine Transform
ANFIS	**:**	Adaptive Neuro-Fuzzy Inference System

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

AUTOMATIC FEATURE EXTRACTION OF EEG SIGNALS USING NEURAL NETWORKS AND TIME-FREQUENCY ANALYSIS

ALSHEAR, Omar Saadi Fathi
Master Thesis, Department of Electrical and Electronics Engineering
Supervisor: Assist. Prof. Dr. Hassan SHARABATY
December 2017, 48 pages

Electroencephalography (EEG) is an important tool for the diagnosis of many human neurological disorders. The visual EEG analysis is very complex and requires a lot of time to be delivered by doctors. Moreover, the manual diagnosis is differing from one doctor to another depending on the doctor experience and some other factors. Time-frequency analysis & Artificial Neural Networks (ANN) is generally used in the automatic diagnosis of these signals. In this work, we propose to use Discrete Wavelet Transform (DWT) to decompose the EEG signals in preprocessing stage depending on Bonn University (UBonn) Dataset, then to analyze the extracted features using 2 hidden layers of ANN in order to deliver the classification decision. The preprocessing analysis was achieved via MATLAB Wavelet Toolbox using two wavelet functions: Daubechies 4 (db4) and Daubechies 8 (db8). In order to study the impact of selected features on the classification accuracy, we have discussed 6 different cases in Feature extraction stage. Finally, the resulted vectors are trained by Levenberg-Marquardt training algorithm in the Decision-Making stage. The performance of our algorithm is evaluated via Confusion Matrix equations. In case of db4, the accuracy of the best proposed scenarios is 98.50% whereas db8 gives an accuracy of 98.75%.

Keywords: Electroencephalogram, Artificial Neural Networks, Discrete Wavelet Transforms, Feature Extraction.

ÖZET

SİNİR AĞLARI VE ZAMAN-FREKANS ANALİZİ KULLANILARAK EEG SİNYALLERİNİN OTOMATİK ÖZELLİK ÇIKARIMI

ALSHEAR, Omar Saadi Fathi
Yüksek Lisans Tezi, Elektrik-Elektronik Anabilim Dalı
Tez Danışmanı: Yrd. Doç. Dr. Hassan SHARABATY
Aralık 2017, 48 sayfa

Elektroansefalografi (EEG), insanlarda bulunan birçok nörolojik bozukluğun teşhisinde kullanılan önemli bir araçtır. Görsel EEG analizi oldukça karmaşıktır ve doktorlar tarafından teslim edilmesi çok zaman gerektirmektedir. Bunun yanı sıra, manuel teşhis doktorun deneyimine ve bazı diğer faktörlere bağlı olarak bir doktordan diğer doktora farklılık göstermektedir. Zaman-frekans analizi ve Yapay Sinir Ağı (ANN) genellikle bu sinyallerin otomatik teşhisinde kullanılmaktadır. Bu çalışmada, Bonn Üniversitesi (UBonn) veri setine dayanılarak önişleme aşamasında EEG sinyallerinin ayrıştırılması için Kesikli Dalgacık Dönüşümü (DWT) kullanılmasını ve daha sonra sınıflandırma kararını vermek için yapay sinir ağının (ANN) 2 gizli katmanını kullanarak çıkarılmış özellikleri analiz edilmesini önermekteyiz. Önişleme analizi, Daubechies 4 (db4) ve Daubechies 8 (db8) olmak üzere iki dalgacık fonksiyonu kullanan MATLAB Dalgacık Araç Çubuğu üzerinden gerçekleştirilmiştir. Seçilen özelliklerin sınıflandırma doğruluğu üzerindeki etkisini incelemek için, özellik çıkarımı aşamasında 6 farklı durumu tartıştık. Nihai olarak, sonuçta ortaya çıkan vektörler Karar-Alma aşamasında Levenberg-Marquardt eğitim algoritması tarafından eğitilmiştir. Algoritmamızın performansı Karışıklık Matrisi eşitlikleri üzerinden değerlendirilmiştir. Db4 durumunda, önerilen en iyi senaryoların doğruluğu %98.50 olurken, db8 ise %98.75'lik bir doğruluk vermektedir.

Anahtar kelimeler: Elektroansefalogram, Yapay Sinir Ağları, Kesikli Dalgacık Dönüşümü, Özellik Çıkarımı.

CHAPTER ONE

BACKGROUND AND LITERATURE REVIEW

1.1 Introduction

The analysis of EEG signals is very complex because of the non-stationary nature of the brain activity. EEG Feature extraction helps to acquire all important and distinctive characteristics stored in the brain signal, and can be used to diagnose various neurological diseases. In addition, it reduces the needed time by doctors to analyze brain signals manually. Many Time-Frequency methods have been proposed to depict the brain wave changes that occur in EEG signals.

1.2 The Brain and Its Structure

The human brain has the same general structure as the brain of other mammals. However, it has evolved to become the largest in terms of size relative to the body. The blue whale has the heaviest brain at 6.92 kilograms compared to 1.5 kilograms for the human brain; however, the human brain has the highest encephalopathy coefficient at seven times higher than the average of mammals [1]. The increase in human cerebral volume comes largely from the development of the cerebral cortex, which is distinct from that of other primates, especially the frontal lobes which account for more than 30% of the cerebral surface and are mainly involved in planning, language and voluntary movement. Nearly half of the cerebral cortex is devoted to sensory analysis, particularly vision.

Although it is protected by the blood-brain barrier and the thick bones of the skull as well as being bathed in the cerebrospinal fluid, the human brain remains subject to injuries and disease, the most frequent being head trauma, neurotoxic diseases, and neurological and neurodegenerative disorders. A number of psychiatric disorders, such as schizophrenia and depression, are considered to be associated with brain dysfunctions, although the nature of these brain abnormalities is not well understood. Figure 1.1 shows the Brain and Human Skull.

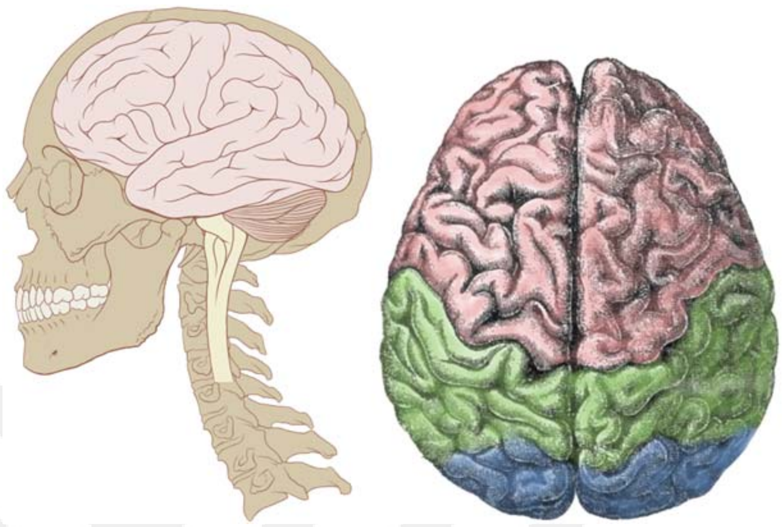

Figure 1.1: Brain and human skull.

Historically, opinions have often been opposed to knowing that the brain or heart was the seat of the soul. In a sense, it was impossible to deny that consciousness seems to be localized in the head and that a blow to the head causes much more unconsciousness than a blow to the torso and that shaking the head gives vertigo. In another sense, the brain subjected to superficial examination seems inert, while the heart beats constantly. Stopping the heartbeat causes death, while emotions induced by changes in heart rhythm and grief often produce a feeling of pain in the heart ("broken heart"). For Aristotle, the seat of the soul was the heart and the brain, a cooling organ, merely served to ensure the circulation of blood (philosophical and medical theory of cardio centrism). Democritus divides the soul into three parts: the intellect in the head, the emotion in the heart, and the desire around the liver [2]. Hippocrates was certain that the soul was in the brain. For Herophilus, the brain was the center of intelligence.

1.3 Literature Review

An epilepsy seizure is a complex illness and it is the second most predominant complex issue in humans after strokes. Between (40) and (50) million people on the planet experience the ill effects of epilepsy [3]. In an epilepsy seizure, the typical example of complex neuronal activities is plainly aggravated, bringing about abnormal sensations, feelings, and conduct or a number of shakes, muscle seizures, and the absence of cognizance. An epilepsy seizure is portrayed as a repetitive seizure by utilizing strange electrical activity created in the brain that adjusts observation or conduct. Epilepsy subjects encounter shifted indications amid seizures relying upon the area and degree of the influenced brain tissue. Contingent upon the degree of the association of brain regions during a seizure, epilepsy can be separated into two fundamental types. Summed up seizures include nearly the whole brain, while incomplete seizures start from an encircled zone of the cerebrum and stay limited to this territory. An electroencephalogram (EEG) is a recording of the electrical activity in the brain. There are two distinct types of EEG which rely upon the terminal regions of the head, namely the scalp (the skin covering the head) and within the skull. For scalp EEGs, anodes are set on the scalp with great mechanical and electrical connections. In spite of this, an intracranial EEG is acquired through exceptional terminals embedded into the brain during surgery. Scalp EEGs that concentration on this exploration are a very widely recognized demonstrative technique to distinguish anomalies of the electrical movement of the cerebrum. EEG records contain important data such as the identification of epilepsy. Inquiries about program seizure location started in the 1970s and different strategies tending to this issue have been exhibited. Chen et al. (2016) proposed a system to utilize DWT and Support Vector Machine (SVM) for epileptic center localization depending on EEG and for giving a rule in choosing the best settings for DWT, disintegrated the EEG windows in 7 commonly utilized wavelet families to their greatest hypothetical levels [4]. Riaz, F., Hassan et al. (2016) presented a system for the discovery of epilepsy seizures in EEGs. Their system depended on the extraction of temporal and spectral characteristics from the Empirical Mode Decomposition (EMD) of EEGs [5]. Singh et al. (2015) utilized wavelet transforms to depict EEG motion in estimated and details coefficients. Spike associated characteristics were extricated for a preparation of counterfeit normal system, which was utilized for characterization of normal patterns and epilepsy patterns in EEGs [6]. Abualsaud et al. (2015) studied the use of a new ensemble of classification to identify an epileptic seizure in the state of compressed and vociferous EEGs. The Noise-Aware Signal Combination (NSC) ensemble classification combines four grouping models in their single execution. The main goal of the proposed classification was to improve the exactness of order within noisy and incomplete information while protecting a sensible measure of many-sided qualities [7]. Kumar et al. (2014) presented a discrete wavelet transforms (DWT) analysis and an approximation entropy (ApEn) of EEGs Seizure discovery was performed in two steps. In the first step, EEG signals were analyzed by DWT to estimate approximation and detail coefficients. In the second step, ApEn weights of the approximation and detail coefficients were estimated [8]. Nanthini et al. (2015) examined the execution of classification concerning Support Vector Machines (SVM) and Artificial Neural Networks (ANNs) by applying wavelet transforms for epilepsy seizure display. The study used the Discrete Wavelet Transform to analyze EEGs that which was non-stationary [9]. Nunes et al. (2014) presented a regular offering valuation of the lately entered Optimum-Path Forest (OPF) classification when working with the assignment of an epilepsy illness diagnosis straight in EEGs. The design used a huge benchmark dataset including five groups; the whole difference was very difficult obtain. The four models from the wavelet transform function and three famous filter methods were examined for the feature extraction and determination, respectively. Furthermore, Support Vector Machines (SVM) composed of the Radial Basis Function (SVM-RBF) kernel, Multilayer Perceptron Neural Networks (ANN-MLP) and Bayesian classification were applied to compare with indications of efficiency and implementation [10]. Fathima et al. (2013) presented a method depending on the wavelet analysis with the computation of certain statistical characteristics. The wavelet analysis performed up to the fourth level, followed by a computation of the statistical features, namely the inter-quartile range of the wavelet coefficients. The features were excerpted for five kinds of EEG signals. A linear classifier trained on these features could classify normal and epilepsy EEG signals with 100% sensitivity, and specificity had an accuracy of 95.6% for five states [11]. Omerhodzic et al. (2013) presented Discrete Wavelet Transforms by t using Multi-Resolution Analysis (MRA) applied for analysis of EEG determination levels of ingredients of the EEGs. Parseval's Theory was applied to obtain the rate of energy distribution characteristics of the EEGs at various decision levels. Second, for artificial neural networks (ANNs), the Features groups of the EEGs depended on the rate distribution of the strength of the signal features [12].

Übeyli et al. (2009) presented a study in which EEGs were analyzed into time-frequency simulations utilizing discrete wavelet transforms with the Daubechies function (order 2) and the analytical characteristics were determined to describe their distribution as supplies to the mixed neural network model [13]. Subasi et al. (2007) presented the discrete wavelet transforms which was used to process and analyze the signal to select the features of the extraction features and followed by presenting ME (Mixture Of Experts), a classification for epileptic seizure discovery dependent on a mixture of expert types [14]. Mohseni et al. (2006) utilized a Short Time Fourier Transform (STFT) to decompose EEG signals and extract properties depending on the imaginary using the Wagner-Ville and pseudo-Wagner-Ville distributions. These characteristics were used as data for artificial neural network classifications [15]. The EEG signals essentially have a small boundary of amplitude (nearly 100 μV) and a spectrum of frequency from 0.4 Hz to 80 Hz. Each EEG was regularly analyzed into five subbands: Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8- 13 Hz), Beta (13-30 Hz), and Gamma (30-60 Hz) [16-18]. Table 1.1 shows the frequency range and capacity for each model of brain waves.

Table 1.1: Spectrum frequency for human brain waves.

Brainwave	Founder	Spectrum (HZ)	Description
Delta	Walter (1963)	0.5-4	Deep sleep
Theta	Walter and dovey (1944)	4-8	Deep meditation
Alpha	Berger (1929)	8-13	Relaxing, without attention
Beta	Berger (1929)	13-30	Solving concrete problems
Gamma	Jasper and Andrew (1936)	Above 30	Sensory processing and solving high cognitive tasks

In [19], the wavelet transform was used to analyze EEG signals to evaluate the mental behavior of human brains. In this paper, the seven emotional states were specified as relax, thought, memory-related, motor action, pleasant, fear, and enjoying music. The proposed algorithm is shown in Figure 1.2.

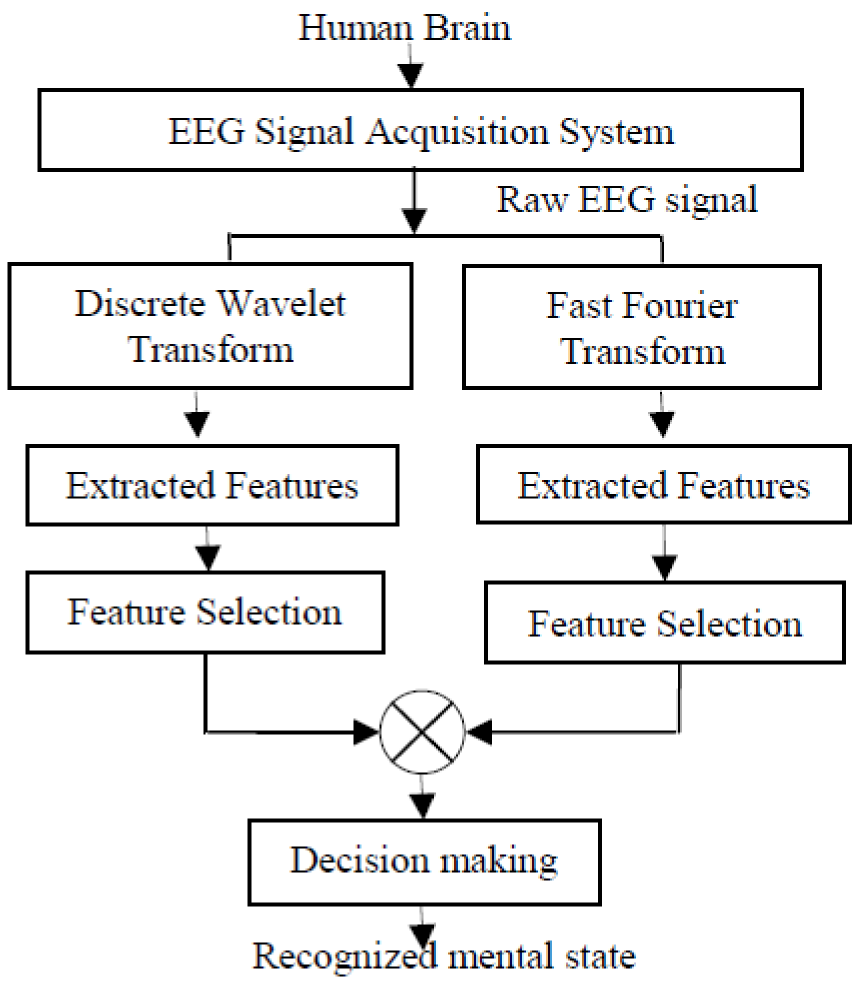

Figure 1.2: Flowchart of the method for [19].

1.4 Problem Definition

The general problem in EEG signals analysis is that the recorded signals have noise in their frequencies. So, the manual analysis is very difficult and require a lot of time to be done by the human. Moreover, the delivered diagnosis is different from one doctor to another [20]. As explored above, many studies have been achieved in order to build algorithms that analyze EEG signal automatically [4, 9, 21]. Generally, time-frequency methods, like DWT with various wavelet functions, are used in the preprocessing phase [8, 9, 11, 22]. However, in the post-processing phase, Artificial Neural Networks (ANN) or Support Vector Machine (SVM) are generally used [10, 12, 15]. Nevertheless, the results of all proposed classification algorithms have shown a margin of error ratio depending on the number of the extracted features and the used wavelet function.

1.5 Proposed Work & Aim of Study

In this thesis, we use Bonn university Database (UBonn) provided by the Andrzejak and his equip in 2001 for different states of brain function. Our work is achieved via MATLAB environments. We propose to use Discrete Wavelet Transform (DWT) with Daubechies order 4 (db4) and Daubechies order 8 (db8) to decompose EEG signal into five sub-bands: Delta, Theta, Alpha, Beta, and Gamma. Six different classification cases will be discussed after extracting 8 features of the EEG signal in both time and frequency domain. In the first case, we will consider the most 4 features used in the previous studies that will be referred as essential features. The remaining 4 features will be considered in the others classification scenarios by combining them with the 4 essential one in order to study the influence of each feature on the classification accuracy. After that, the resulted vectors in all cases is given to the Artificial Neural Network (ANN) which will be trained by Levenberg-Marquardt. Finally, we evaluate the performance of all cases using the Confusion Matrix equations.

This work aims to obtain high classification accuracy between healthy and non-healthy EEG signals.

1.6 Thesis Organization

In Chapter 2, we present the feature extraction stage with explaining to different DWT techniques used to analyze the EEG signals. Whereas the classification with the Artificial Neural Network is included in Chapter 3. After that, we present and discuss our results in Chapter 4. Finally in Chapter 5, we summarize all conclusions of our work with proposed idea for the perspective works.

CHAPER TWO

FEATURE EXTRACTION & WAVELET TRANSFORM

2.1 Introduction

The wavelet transform has two major differences from the Fourier transform in the short term: It can implement a different basis, not necessarily sinusoidal; there is a relationship between the width of the envelope and the frequency of oscillations, thus not only does it perform a scaling of the oscillation, it performs a scaling of the wavelet.

However, it is not a different formalism of the Fourier transform; it is, in fact, complementary. It is wavelet decomposition using Fourier formalism. The wavelet technique is particularly used in computer data compression [23].

2.2 History of Wavelet Decomposition

Wavelets emerged when some subjects of study required an analysis of frequency and time. In the nineteenth century, Fourier analysis was the only technique for the decomposition of a signal and its reconstruction without loss of information. Unfortunately, it provides a frequency analysis but does not allow the temporal location of abrupt changes, such as the appearance of a second musical note after a first note is played.

In 1909, Alfréd Haar defines a function composed of a short negative pulse followed by a short positive pulse as the first wavelet [24].

In 1946, Hungarian mathematician Dennis Gabor invented the transformation of a function similar to that of Joseph Fourier, applied to a time window expressed by a Gaussian function. Finally, the term 'wavelet' was introduced into the mathematical language by Jean Morlet and Alex Grossmann in 1984. The term, originally French, was translated into English as ‘wavelet,’ from the terms wave (wave) and the suffix ‘-let,’ denoting a diminutive (small).

Yves Meyer, recognized as one of the founders of wavelet theory, gathered in 1986 all previous discoveries (numbering 16) then defines orthogonal wavelets [25].

In the same year, Stéphane Mallat [26] made the connection between wavelets and multiresolution analysis. Finally, Ingrid Daubechies devised, in 1987, orthogonal wavelets called Daubechies wavelets that are easily implementable and used in the JPEG 2000 standard [23, 27].

2.3 Mathematical Definition

There are two types of wavelet subgroup: *discrete* and *continuous*.

2.3.1 Continuous Wavelet Transform

Analyzing a square sum able function wavelet is to calculate all its scalar products with the wavelet family. Numbers obtained are called wavelet coefficients, and combining the transaction into a function of its wavelet coefficients is called a wavelet transform [23].

2.3.2 Discrete Wavelet Transform

We can adapt the wavelet transform in cases where it is in a discrete set. This technique is used in the compression of digital data, with or without loss [23].

2.3.3 Using Wavelet Decomposition

Wavelet decomposition is used particularly in data compression. The image compression method is used mainly in two formats: Enhanced Compression Wavelet (ECW), used primarily by professional mapping; JPEG 2000, from the newest ISO standardized format.

This compression method is also used for video: The Dirac codec without a patent allows resolutions from 176×144 (QCIF) to 1920×1080 (HDTV), progressive or interlaced, double compression and better quality (almost lossless) in comparison to MPEG2.

It is based on the wavelet used for the compression by eliminating non-perceptible high-frequency information by the eye. In particular, this often allows better analysis of functions with discontinuities or local phenomena. This is, for example, if the contours in the images, which explains the adoption of a wavelet decomposition in the JPEG 2000 standard [23].

One-dimensional (1D) signals, such as voltage signals, can be analyzed by using wavelet transformation. It is well known that wavelet transformations are widely used to compress data and to suppress noise. In the literature, this process is known as wavelet de-noising.

Wavelet transformations of (1D) signals contain information regarding any low-frequency content of the voltage signal. This low-frequency content is called the *approximation* coefficients, while the information regarding the high-frequency content of the voltage signal is called the *detail* coefficients.

The wavelet transformation of the signal using the matrix method is described in the following sections.

2.4 Wavelet Transform

Let $\Psi(t) \in L^2(R)$ is a continuous-time mother wavelet function and the set of functions obtained by shifting and scaling the mother wavelets:

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \tag{2.1}$$

is the orthonormal wavelet basis in the $L^2(R)$. That is:

$$\int \psi_{a,b}(t)\widetilde{\psi}_{a',b'}(t)dt = \delta(a-a')\delta(b-b') \tag{2.2}$$

In the wavelet, the *a* and *b* variables are real and the integral value indicates the closeness of the signal to a particular basis function. Dividing $\Psi_{a,b}(t)$ by $\sqrt{a}$ ensures the unity in the L^2 norm of the set $\{\Psi_{a,b}(t)\}$ [28, 29].

The main disadvantages of the CWT are computational complexity and redundancy. The mother wavelet has to satisfy the following properties [30].

A wavelet must have finite energy

$$\int \left|\Psi(t)\right|^2 dt < \infty \tag{2.3}$$

The time-scale cells corresponding to dyadic sampling are shown in Figure 2.1.

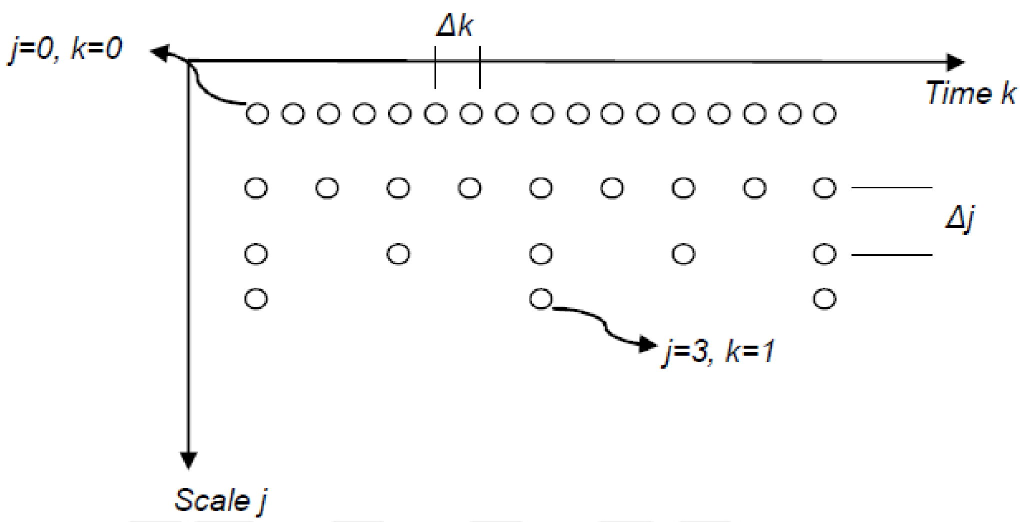

Figure 2.1: Time-scale cells corresponding to dyadic sampling [31].

The discrete wavelet which is generated by dyadic sampling from the continuous wavelet transform is given by:

$$\psi_{j,k} = 2^{j/2}\psi(2^j t - k) \tag{2.4}$$

$\Psi_{j,k}$ is known as wavelet basis and we employ linear combinations of basic functions which are localized both in time and frequency to construct a signal function *(f(t))* which is a linear combination of basic functions. Therefore, we can express a signal function as [32]:

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{j,k}(t) \tag{2.5}$$

where

$$b_{j,k} = \int_{-\infty}^{\infty} f(t)\psi_{j,k}(t)dt \tag{2.6}$$

2.5 Multi-Resolution Analysis

A Wavelet Transform leads to a signal decomposition technique known as multiresolution analysis (MRA), which analyzes the signal at different frequencies at different resolutions.

The idea is that there is a scaling transformation which moves in discrete steps, up and down an associated scale of subspaces. One of the resolution scales refers to "coarse" and the other to "fine." When we compare two subspaces, we can see that the space of the coarse scale is contained within that of the fine resolution [33].

Where *j* and *k* are the dilation (scale) and translation indices, respectively, and can take on only integer values. Dilation and time parameters determine frequency and the time resolution of the WT.

A function ($f(t)$) in the whole space has a piece in each subspace. Those pieces contain more of the complete information in $f(t)$. The piece in V_j is $f_i(t)$. One requirement on the sequence of subspaces is completeness:

$$f_i(t) \rightarrow f(t) \quad as \quad j \rightarrow \infty \tag{2.7}$$

We let W_0 be space spanned by the orthonormal set of basis $\{\psi(t-k), k \in N\}$. Space W_0 is orthogonal to the space W_1. Then spaces spanned by the wavelet function bases are orthogonal between each other. Thus, $\ldots W_{-1} \perp W_0 \perp W_1 \perp W_2 \perp \ldots$

Any signal in the V_1 space may be expressed in terms of the basis of V_0 and W_0 space. If we combine the bases of V_0 and W_0 space, we can define any signal in V_1 space as:

$$V_1 = V_0 \oplus W_0 \tag{2.8}$$

W_0 is the complementary of V_0 while V_0 is a subset of V_1. Therefore, the V_0 and W_0 spaces are complementary. Two spaces are called that satisfy this property of orthogonality and use V_j ^W_j to denote V_j as orthogonal to W_j. Their bases together can represent any signal in the next "higher" or finer space of V_1 [34]. V_j denotes subspaces corresponding to scaling basis (approximations) and W_j denotes subspaces corresponding to the wavelet basis (details).

The part of the signal at resolution *j* and space V_j is the approximation of the signal at this resolution:

$$a_j(t) = \sum_{k=-\infty}^{\infty} a_k \phi_{j,k}(t) \tag{2.9}$$

and the part of the signal at resolution j and space W_j is the detail of the signal at this resolution:

$$d_j(t) = \sum_{k=-\infty}^{\infty} \beta_k \psi_{j,k}(t) \tag{2.10}$$

The signal at resolution j is then:

$$f_j(t) = a_j(t) + d_j(t) \tag{2.11}$$

The relationship between the scaling and wavelet function spaces is shown in Figure 2.2. Spaces spanned by scaling function bases are nested. Each V_j is contained in the next subspace V_{j+1}.

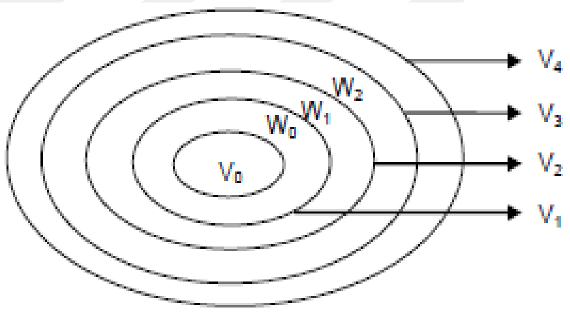

Figure 2.2: Nested vector spaces spanned by scaling and wavelet bases.

The wavelet decomposition process is represented schematically in Figure 2.3.

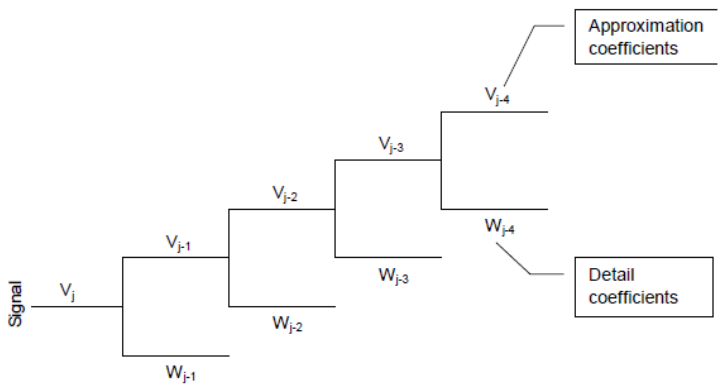

Figure 2.3: Schematic representation of wavelet decomposition of signals.

Below the properties, the multiresolution analysis is summarized.

CHAPTER THREE

CLASSIFICATION VIA ARTIFICIAL NEURAL NETWORK

3.1 Background & Definition

Computers have become an indispensable part of the modern world and computer systems today can decide on two events and learn about the relationship between those two events. This cannot be expressed mathematically and it cannot solve problems that can be solved by computers using methods based on experience. Computers equipped with this feature, and leading the development of these capabilities work "Artificial Intelligence," is known to be working. While working on artificial intelligence, scientists have been forced to work on the brain and being inspired by the structure of the brain, they have tried to extract a mathematical model. In order to model the brain with the idea that all behavior should be modeled accurate, the physical components of a variety of artificial cell and network models have been developed. Thus, with today's computers as a separate discipline of algorithmic calculation methods, the "neural network" has emerged. In the most general sense of the ANN, it can mimic many nerve cells in the human brain that are the result of simple processors connected together at different levels of impact regarded as a complex system. Developed as a result of efforts to remove the mathematical model of nerve cells in the human brain, the ANN, in practice, often has a very different structure and form of data which may be quickly identified and used for detection. Engineering ANN is a reason behind the widespread use in the field, the ability of the solutions put forward in the classification technique with an effective solution to difficult or even impossible problems. Computers are very successful in their mathematical and algorithmic calculations in the process requiring a great amount of learning and recognition systems with the expectation of accuracy to perform operations. This matter is best resolved by an ANN. Neural networks can be used in nonlinear equations and generalization capabilities due to the very complex and large-scale solutions to problems that are easily produced. Adaptable to be fast, easy, and in order to analyze and design as well as due to their ability to learn, they have become one of the most indispensable elements of our age [35].

Results that can be obtained with a very complex equation using other programs were made possible through the foresight of the neural network that can be estimated easily. Machines that can learn using neural networks due to in mind that because of the way people work areas or constraints of the human factor in the work to be done in the coming loss of precision has been achieved major improvements occur. A parallel ANN can be defined as a distributed computing system; in other words, on the basis of the ANN, they are formed in the process room functions requiring intelligence. This system consists of a one-way beacon channel and interconnected operation elements. The output signal can be amplified by the fact that although the cases need one, externally it is a simple though there appears to be an extremely complicated internal structure in the ANN. Because it is an imitation of biological neural networks, an understanding of the structure of biological neural networks will facilitate the understanding of the neural network. The building blocks required in artificial neural networks are neurons as in biological neural networks [35, 36].

Containing approximately 10 billion neurons in the human brain, the nerve cells are estimated to comprise 60 trillion connections with each other. Nerve messages from the senses are transmitted to the next cell evaluated by nerve cells. Thus, the message carried by the signal is transmitted to the central and nervous system. This generates response signals to evaluate central nervous system signals. These signals are transmitted in response to the organ in the form of the nervous system. Thus, the signals from the response of the sensory organs will be transmitted to the body through the nervous system in spite of the much slower process of the brain connected by parallel connections and experience gained in this enormous ability to renew connections between neurons are the ANN application directory [37].

At this point, the general structure of the ANN will be useful to sort out before discussing the details about the features that make it important, including the most basic capabilities of ANN in this context.

3.1.1 Non-Linear Function of The Network

The neuron itself is not linear but it is the most essential component in an ANN. Accordingly, the neural network formed by the interconnection of neurons is not linear and also the non-linear characteristic due to the nature of an ANN is distributed across the network in parallel. This is mathematically possible in the absence of the desired linear mappings due to non-linear sub-units on the network structure in the fulfillment of the function. Therefore, the function of structural flexibility allows this to occur correctly.

3.1.2 Input-Output Association

The ANN guided especially to be mentioned later in the study has a structure that allows learning to occur so that using the elements in the ANN training set provide the desired results for the specified entry revising free elements in its structure. To make fewer errors at each iteration, it resets the weight and connections.

3.1.3 Becoming Acceptable Easy Adapter

When the threshold value is changed or by simple operations such as adjusting, the weight applied to the input can be adapted to the new environment. ANNs with different structures are used in different fields and thanks to this ability, they may be more preferable.

3.1.4 Parallelism and Function of Disarray Structural

Neurons within the network structures actually working simultaneously constitute a complex function. Working simultaneously in multiple neurons does not directly affect the success of the network in the event of outages of any neurons. Depending on the application in order to obtain the proper output, connections between neurons can be canceled automatically. This is indicative of the fact that in parallel and distributed structure of all the neurons in achieving the expected results of the ANN.

3.1.5 Generalization Ability

The network during training describes the mapping of numerical data used if using the coarse features and it is thus able to produce meaningful answers for unused inputs during training. The neuron structure of the network in the event of breakdowns is shown in Figure 3.1.

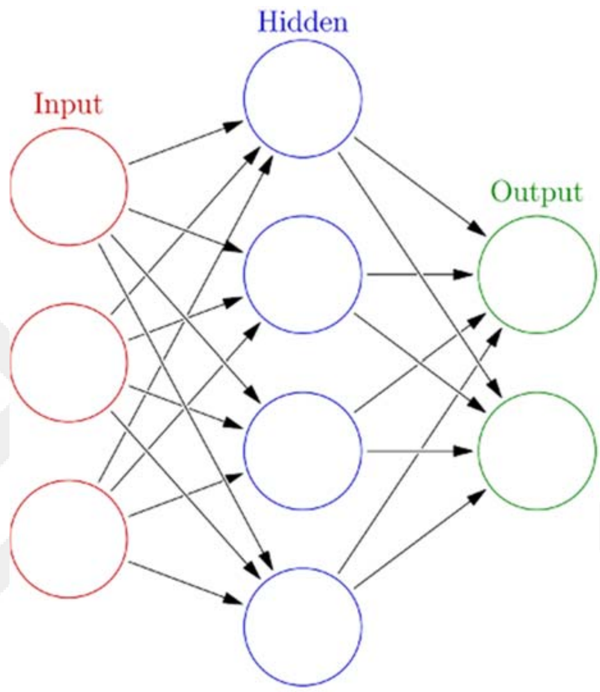

Figure 3.1: Neuron structure of the network in the event of breakdowns [38].

3.1.6 Error Tolerance

An ANN application of a possible error occurring in the system is also one of the most important advantages such that the quality will not allow a sudden collapse. An ANN error will occur in the structure above because of the mentioned features that will result in gradual decline and the user will notice that something is wrong and provide the time needed to measure it.

3.2 General Structure of Neural Network

Neurons, the building blocks of artificial neural networks, meet in many layers that can be used in solving highly complex problems. An ANN has four main parts forming dendrites, axons, the core, and connections, as shown in Figure 3.2.

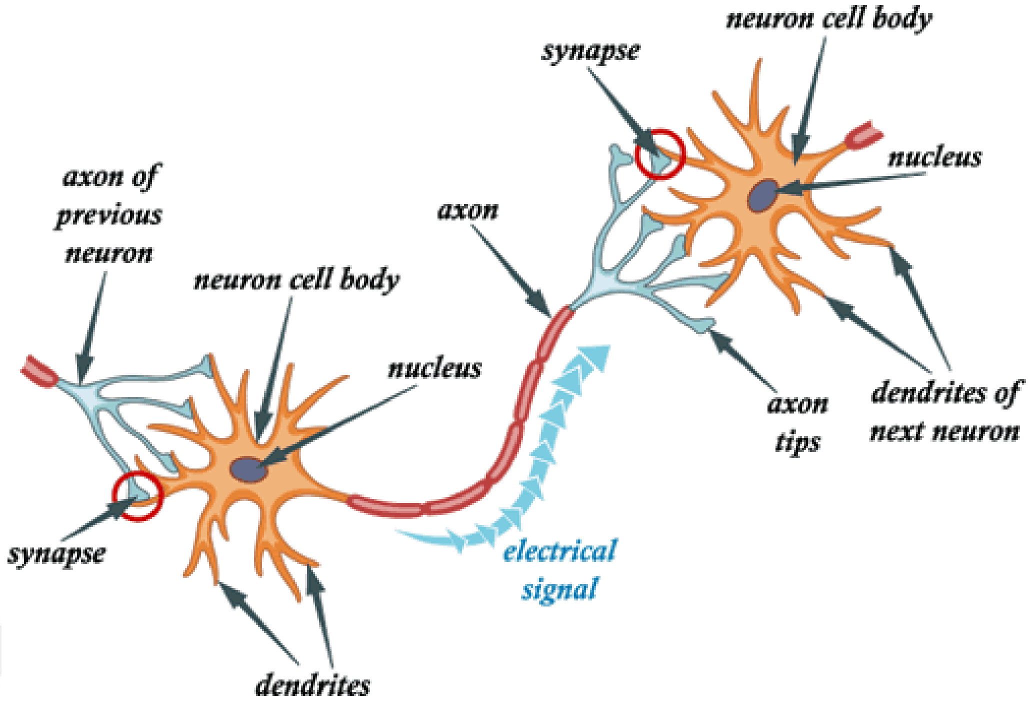

Figure 3.2: Sections comprising a neural network [38].

3.2.1 Dendrites

Dendrites are tree root-like structures which convey the core of the electrochemical excitation received from other neurons. The electrical stimulation of neurons and dendrites in a consecutive sequence reaches through the synapses between them.

3.2.2 Core

The central processing unit core of the nerve cell collects and transmits signals from the dendrites to the axon.

3.2.3 Axon

Show and branching out from the body which is the cytoplasmic fraction. For each neuron; there is one. Axon exiting the body, to move from the nerve cell signals in the environment is official, so other nerve cells, nerve cells, muscle cells, such as cells or operational ties with the gland. Has an important role in message transmission.

3.2.4 Connection

Newly produced is responsible for transmitting signals to other neurons. To receive notifications of a nerve cell is called the threshold stimulus intensity required intensity. The amplitude of the said notice must be at or above the threshold in order to receive alerts. The warning threshold value below the response is not generated by neurons.

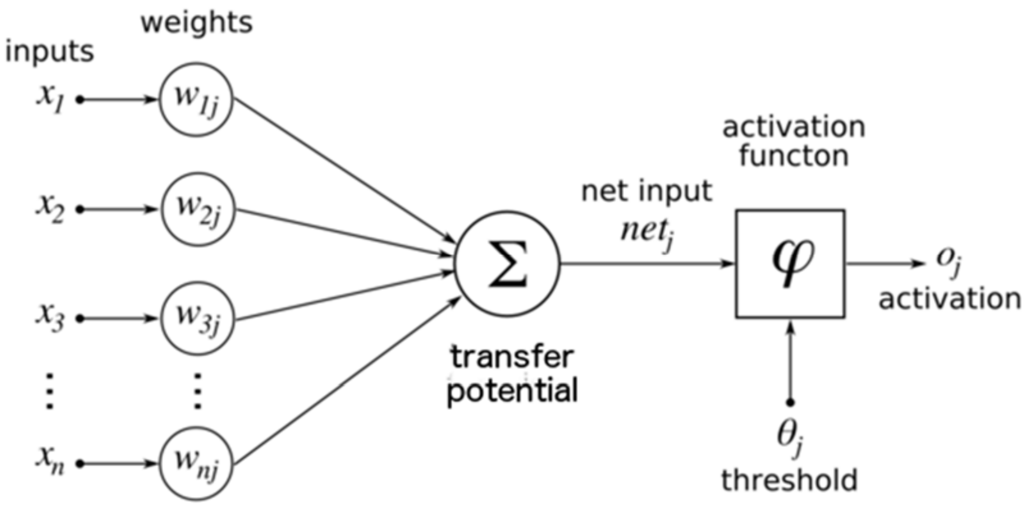

Figure 3.3: Artificial neuron structure [38].

The neuron model, shown in simplified form in Figure 3.3, can also be considered a threshold volume. The threshold (θ) unit, which collects the output and Hazel Harmandir, only produces an output exceeding the sum of the internal threshold of the entrance. A threshold neuron unit receives signals from the synapse and collects all the signals that are generated by multiplying the appropriate weight. You will now sign up to the power of the threshold gathered strong sign stimulus is transmitted along the axon and dendrites of the other neurons. Intersections compared with all the signs of the internal threshold gated neurons with dendrites and synapses from the axon signs spreading the threshold is exceeded. Output depending on whether the result of the sum function value is above or below the threshold activation functions are formed by the normalizing. ANN, connecting these simple nodes and the unit is obtained by conversion to a network. An ANN block diagram is shown in Figure 3.4.

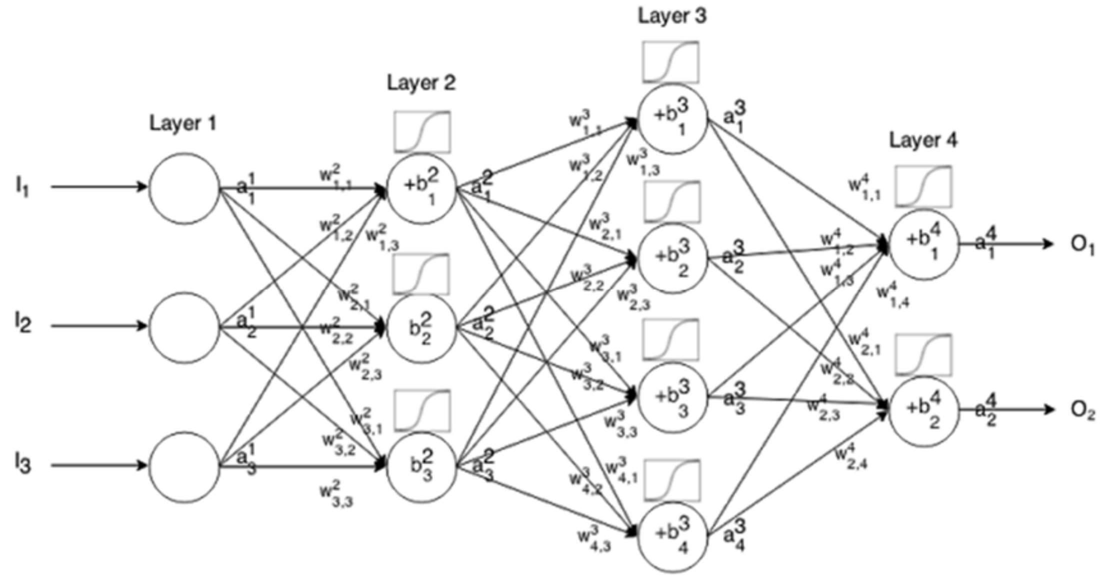

Figure 3.4: ANN general block diagram [38].

To set the threshold value of the signal entering the activation function at this point bias (θ) is called a fixed input value applied to the input neurons, which allows the threshold to be changed in order to receive the expected outputs. An ANN has a basic, simple and versatile structure in which each node cell is called the *n*th degree and a non-linear transfer. Processing elements called nodes and links there between, each connection is involved in the transmission of a one-way mark delay. An unlimited number of processing element inputs and a single output connection. However, if this needs to be copied and used in many cells of single output power supply, the output of the processing element may be any desired mathematical type. The function of the output of the processing element x (the input value), y (the output value), Ψ (the transfer function, $f(x)$ (the collecting function), W_{kj} k (the connection weights) and θ_k for sequential neurons, including neurons in the sequential threshold for k, can be expressed as follows [33]:

$$f(x) = w_{k1}x_1 + w_{k2}x_2 + \ldots + w_{kn}x_{kn} + \theta = \sum_{j=0}^{n} w_{kj}x_j; \quad \theta = w_{k0}x_0 \tag{3.1}$$

$$y = \psi(f(x)) = \psi\left(\sum_{j=1}^{n} w_{kj}x_j\right) \tag{3.2}$$

expressed in the form of a neuron matrix as follows:

$$f_k(x) = [w_{k0}\ w_{k2}\ w_{k3} \ldots w_{kn}] \begin{bmatrix} x_0 \\ x_1 \\ . \\ x_n \end{bmatrix} = w_k^T x \tag{3.3}$$

Inputs are information that enters the cell from other cells or from the external environment. The information enters the cell via links on weight or weights, which will determine the effect on the corresponding input cell. The input values are weighted after entering the processing element; in other words, the impact on the system of each input signal can be replaced by the weight assigned to it. Weights are the relative strength of the value to be used as inputs in a neuron (the mathematical coefficient). Multiplication is used at this point. One neuron usually simultaneously is entered as many numbers. There are different weight values of every connection to the transmission between neurons in the ANN input. Thus, the weights impact on all inputs of each processor element, which has a weight of its own at each entry. These weights have the same function as the varying biological effects of a synaptic neuron. In both cases, some entries will be more effective in producing a neural response as the ensuing process becomes more an important element than the others. Commonly known as the aggregation function, it calculates the input coming from a cell and it is usually entered and calculated as the sum of the product of the entries related to weight.

3.3 Learning in Neural Networks

The network should be adaptable to a good result from artificial neural networks. This is only possible with suitable weights and the right connections. The network connection must learn the proper weight and behavior of the system in order to obtain organizing itself. Information in a neural network is kept in the weights of the neural connections in the network. Therefore, it is important to determine the weight. The knowledge that the weight value is stored in the entire network with a node does not mean anything to a single press. Weights across the entire network should take the appropriate value. However, there is a formula originally developed to determine the values of the optimum weights in a neural network. Network processor elements determine optimal weight values using a set of rules over time.

This processing is called "train the network". Accordingly, the weight value of a network to be trained must be dynamically varied within certain rules [34].

Overall, a learning event in a neural network takes place in two stages. Weight values to be shown in the first stage random sample are taken and the network has to be generated by the output of the network. According to the accuracy of the output value of the second stage as different examples showing the weights of the network are changed. The objective here is to find a weight value which will be useful to obtain an accurate output for these samples. Finding weight values to produce the correct output of the network indicates that it has the ability to make generalizations about the events represented by the example of the network.

The event of generalization of the network is called a "learning network". Learning methods, including supervised, unsupervised and reinforcement were collected into three basic groups [35].

3.3.1 Supervised Learning

In this method, the intervention of an outside trainer should say to the neural network. Trainers should produce results relevant to neural network input data in the neural network system. Therefore, the artificial neural network to input/output samples comprises two presentations.

These two represent the features required to learn the network. This is part of the network input and it generates an output current connection with the information represented by the weight. This output is compared with the output which should be transferred over a network error and intermediate weights are modified to reduce this error.

3.3.2 Unsupervised Learning

The tutorial which helps the learning unsupervised learning network. Therefore, it is also often not called self-organizing (self-organized learning). Networking acquires its examples shown and classifies them according to certain criteria. These criteria can be known in advance. The network itself constitutes their learning criteria.

3.3.3 Reinforcement Learning

This method is close to supervised learning rule learning. The reinforcement learning algorithm does not need to know the desired output. It does not provide an ANN output to the target output, the accuracy of a measure that is used to input the corresponding output obtained.

3.3.4 Back Propagation Algorithm

The back-propagation algorithm is the most widely-used learning algorithm for updating the parameters of a neural network. Today, problems with speech recognition solutions as artificial neural networks to the problem of non-linear systems are used with success in many fields [36].

Due to the attempt to reduce the error in the reverse direction, the network algorithm is called to the back-propagation algorithm. Today, it is derived from the many developed versions of the back-propagation algorithm. However, the back-propagation algorithm is usually expressed with the generalized delta learning algorithm.

Calculation of the back-propagation algorithm consists of two parts:

a) Advanced calculations

b) Back calculation

3.3.5 Advanced Computing

Advanced methods of calculation start with the administration of the network from the input layer of each sample in the training data set. Input data is sent to the intermediate layer between the hidden layer and input without any change. Here, the collection function is applied to each of the neural cells in the hidden layer. Input is gathered in the hidden layer of neural cells, with the result of the threshold value of the collection function being calculated. As in biological neural cells, they are normalized to the net input of the activation functions possessed by every neural cell to generate an electrical signal and producing an output value for the cell. This creates an output value of the input value of the neural cells in the next layer, and continues until the output of the computation network. The exit output value is found in layers and advanced calculation of the network is completed.

3.3.6 Levenberg Marquardt Algorithm

Basically, the Levenberg-Marquardt (LM) calculation is a minimum square estimation calculation subject to the greatest neighborhood outline. Let $E\ (w)$ be a target error work surrendered of one of a kind blunder terms $e\_i\hat{}2\ (w)$ as result:

$$E(w) = \sum_{i=1}^{m} e_i^2(w) = ||w||^2 \tag{3.4}$$

Wherever $e_i^2(w) = (y_{di} - y_i)^2$ and y_{di} are the needed values of the output neurons, y_i is the exact result of that neuron. It is assumed that function $f\ (\bullet)$ and its Jacobian J are verified at point w. The reason for the Levenberg-Marquardt calculation is to ascertain the weight vector w with the end goal that $E(w)$ is a minimum. Using the Levenberg-Marquardt calculation, another weight vector $wk + 1$ can be taken from the past weight vector w_k as:

$$w_{k+1} = w_k + \delta w_k \tag{3.5}$$

where δw_k as:

$$\delta w_k = -(J_k^T f(w_k))\,(J_K^T J_K + \lambda\, I)^{-1} \tag{3.6}$$

In Equation (3.6) J_K is the jacobian of f estimated at w_k, λ the Marquardt parameter, and I the character matrix.

3.4 Statistical Parameters

The estimation of the suggested system on the identification problems is defined by calculating the statistical characteristics parameters of sensitivity accuracy, specificity accuracy, and classification accuracy. These statistical characteristics parameters include:

Sensitivity True Positive Ratio (TPR): the number of accurately identified sure patterns / all numbers are actually sure of the patterns. A sure pattern shows the detection of epilepsy:

$$\text{Sensitivity (TPR)} = \frac{\text{TP}}{\text{TP+FN}} * 100\% \tag{3.7}$$

Specificity True Negative Ratio (TNR): the number of accurately identified unsure patterns / all numbers are actually unsure of the patterns. An unsure pattern shows the detection normal/not-epilepsy:

$$\text{Specificity (TNR)} = \frac{\text{TN}}{\text{TN+FP}} * 100\% \tag{3.8}$$

Classification Accuracy (CA): the number accurately identified models/all numbers exemplars.

CHAPTER FOUR

PROPOSED ALGORITHM, RESULTS AND DISCUSSIONS

4.1 Proposed Algorithm

Our proposed algorithm consists of three sections as shown in Figure 4.1. First of all, the preprocessing stage is achieved by using discrete wavelet transform that decompose the EEG signal into five sub-bands at specific frequencies. Then, eight features are extracted from the EEG signals in both time and frequency domain in multiple scenarios in order to measure the effect of the secondary features. Finally, these features are fed into an Artificial Neural Network to obtain the final decision. This algorithm is achieved using MATLAB 2017a.

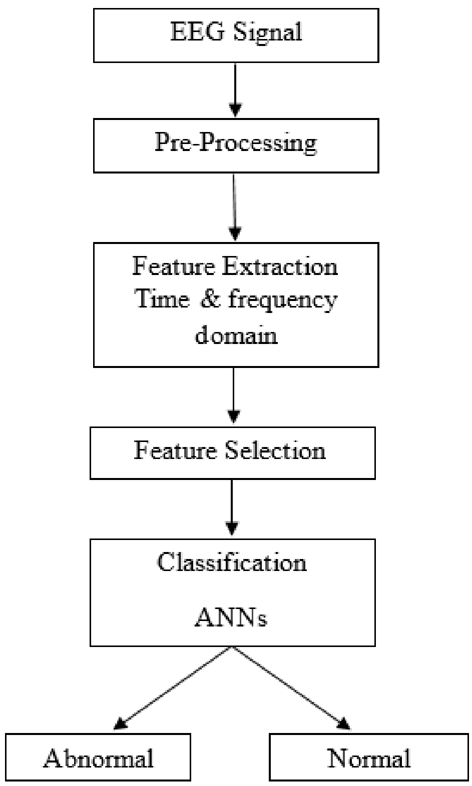

Figure 4.1: Frame general for analysis and classify of EEG signal.

4.2 Subjects and Datasets Discription

The used Dataset in this work was provided by Bonn University for research purposes [39]. The entire information collection comprised five models (named A-E) each including 100 single channel EEG windows. These windows were selected and framed from constant multi-channel EEG records taking after visual examination for artifacts, such as being the result of muscle development or eye developments. Sets A and B comprised windows obtained from outside EEG records that were performed on five ordinary volunteers by applying an institutionalized final plan framework, as shown in Figure 4.2.

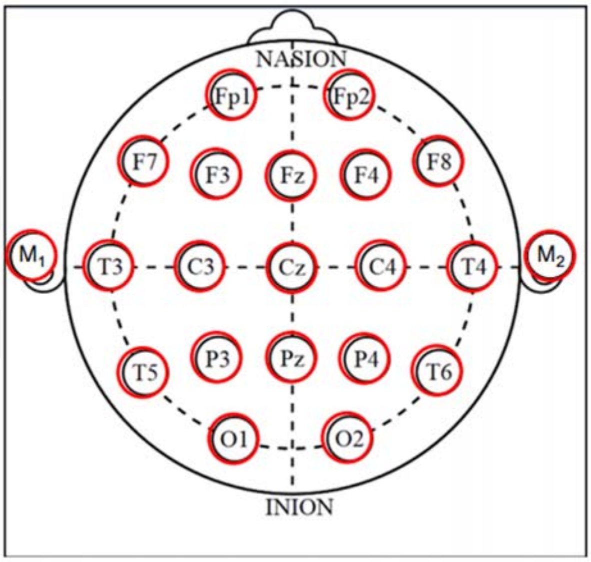

Figure 4.2: 10/20 Global method for arranging the electrodes.

Volunteers were stimulated in a wakeful state with eyes open (A) and eyes shut (B), consecutively. Datasets C, D, and E additionally from EEG document of pre surgical assurance. EEGs from five patients were selected, every one of which had performed full seizure control after resection of one of the hippocampal shapes that were in this way very much analyzed to be the epileptogenic put. Windows in set D were recorded from inside the epileptogenic zone, the area that contained seizure action. The windows were examined over all recording locales indicating octal action. All EEG signs were recorded with the comparing 128-channel enhancer mode, utilizing a normal regular reference. The information was digitized at 173.61 specimens per second utilizing 12-bit determination. Band-pass channel surroundings were 0.53-40 Hz (12 dB/oct) and duration of window 23.6 second. In this study, we focused on the two Datasets (A and D) of the entire Dataset. Standard EEGs are presented in Figure 4.3.

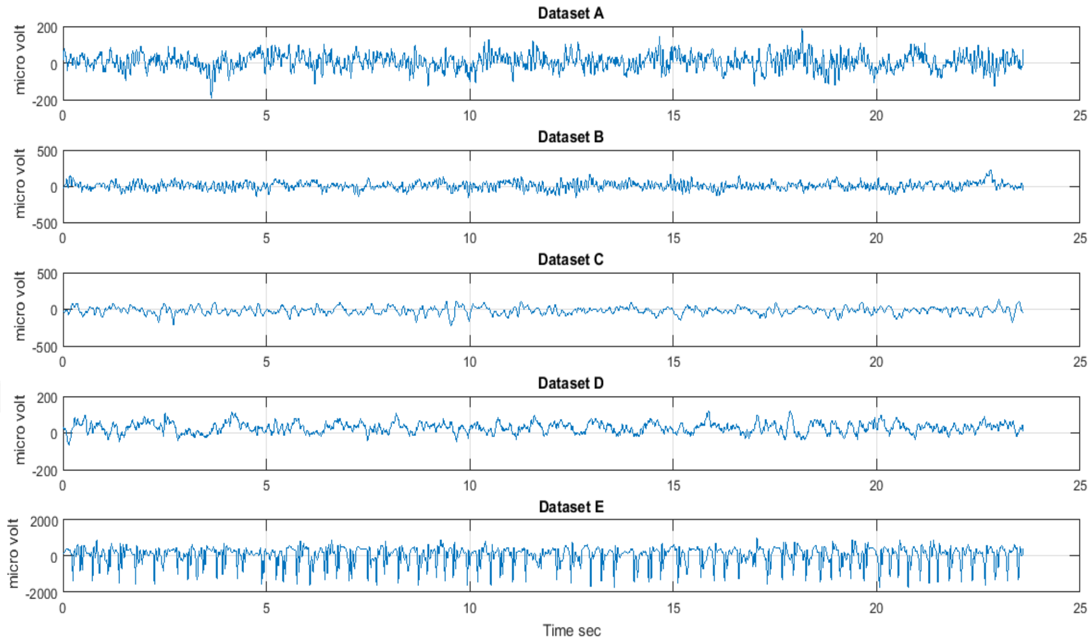

Figure 4.3: Five datasets of UBonn database.

4.3 Analysis Levels

The analysis level is an essential parameter of the DWT. Each level in a DWT agrees to a particular frequency band. Moreover, levels of decomposition give further detailed depictions of signals but they may introduce feature repetition and increase the computational cost. Based on the classifier being used, the computational cost can rise exponentially. The highest-level L of analysis, in theory, is combined decided by the signal and the mother wavelet to complete the condition:

$$L < log_2 \frac{n}{F-1} \tag{4.1}$$

Where n and F are the signal length and filter length, respectively [40]. Any window in the UBonn dataset has 4097 sampling points and each wavelet has a highest theoretical level of analysis provided in the column 'Max Level,' as shown in Table 4.1.

Table 4.1: Max level for Daubechies function on the UBonn dataset.

Wavelet	Max level
db4	**9**
db8	**8**

4.4 Feature Extraction using DWT

EEGs can be considered an overlap of various formations occurring and approaching various time-scales at various intervals. For one of the concepts, wavelets decompose to separate and identify underlying configurations at various time scales. This means that we know that the Wavelet Transform (WT) fits to explain non-stationary signals because it is extremely limited in the time-frequency domain. Each part from the time and frequency identification implies recognition as compact support which is one of the most interesting characteristics in the WT. The central influence is in the WT because the WT has to change windowing size, which is wide at lower frequencies and small at the high frequencies, thereby driving an improvement in the time-frequency resolution in all frequency spectra. Hence, the spectral decomposition of the EEGs occurs by utilizing the DWT. The choice of the appropriate wavelet and the number of analysis levels is actually influential in the analysis of signals utilizing the WT. The number of decay levels is based on the predominant frequency elements of the signal. The decomposition levels are arranged where the signal bands that correlate excellently with the frequencies needed for distribution of the signal are maintained in the wavelet coefficients.

In this study, the number of levels was selected to be (5). Therefore, the EEGs were analyzed into details D1–D5, the last being an approximation, A5 as shown in Figure 4.4 and Table 4.2.

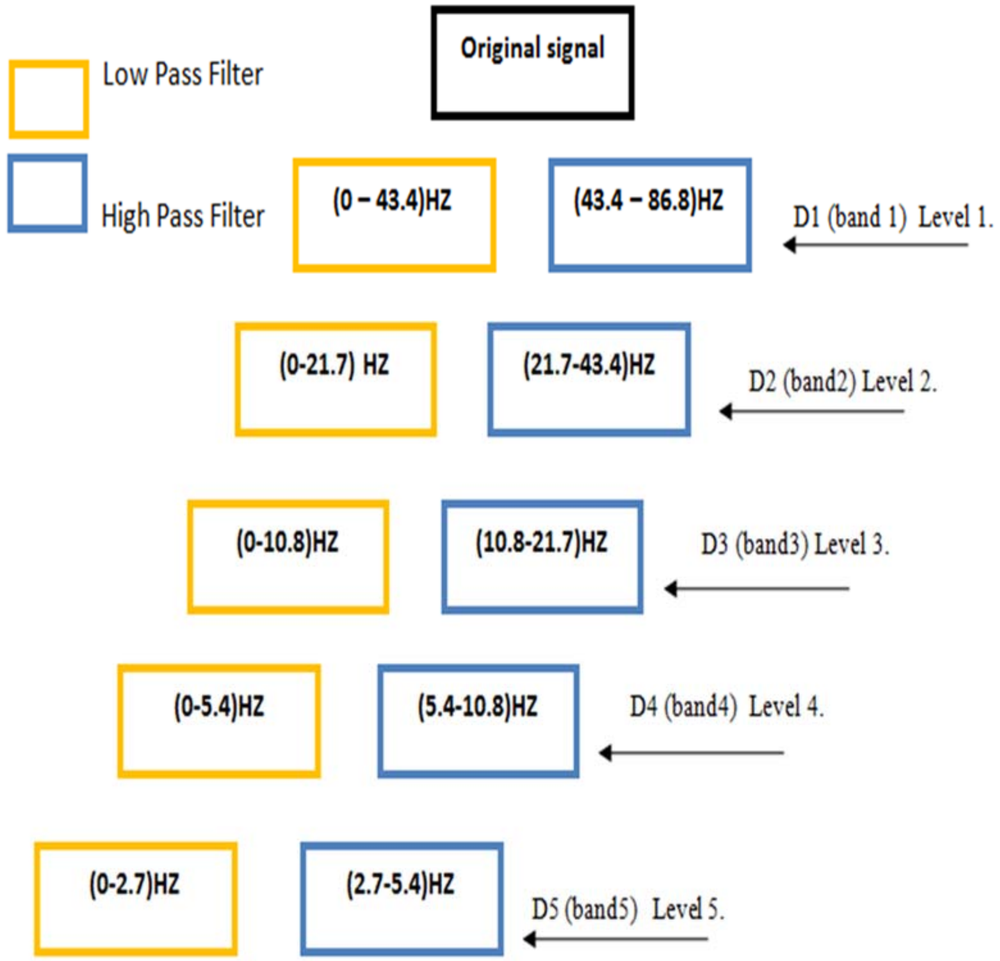

Figure 4.4: EEG sub-band analysis. 5-level analysis (UBonn Dataset).

Table 4.2: Bandwidth of frequencies for all coefficients.

Decomposed Signal	Frequency Range (Hz)
D1 - Non-useful frequency (noise)	43.4-86.8
D2 - (Gamma)	21.7-43.4
D3 - (Beta)	10.8-21.7
D4 - (Alpha)	5.4-10.8
D5 - (Theta)	2.7-5.4
A5 - (Delta)	0-2.7

Normally, experiments are conducted with various kinds of wavelet function and the one that provides the best performance is selected for the appropriate application. The similarity features are found in the Daubechies wavelet and it is more appropriate to identify changes of the EEGs. Hence, the wavelet coefficients are calculated employing the Daubechies wavelet function of order (4) and Daubechies order (8). In this study, the wavelet transforms coefficients were calculated using MATLAB software. The ANN inputs represented the most significant element to create the artificial neural network dependent on the model classifier as the best classification will perform badly if the information is not chosen well. This input determination has two purposes: (1) they are the components of a model, or (2) they are established inputs that best serve a presented pattern. The estimation of the discrete wavelet coefficients gives a compact design that displays the energy spread of the signal in the time-frequency.

Hence, these detail wavelet coefficients and approximation wavelet coefficients of the EEGs were applied to the vectors of those features expressing the information in the signal. Rectangular windows were composed by 4097 discrete data and they were selected so that it contained a single EEG window. For each EEG window, the details of the wavelet coefficients were calculated (DK, k = D1, D2, D3, D4, D5) at the 1st, 2nd, 3rd, 4th and 5th levels (2052 + 1029 + 518 + 262+134 wavelet coefficients), and one approximation wavelet coefficient (A5) at the 5th level (134 wavelet coefficients) was calculated, as shown in Figure 4.5.

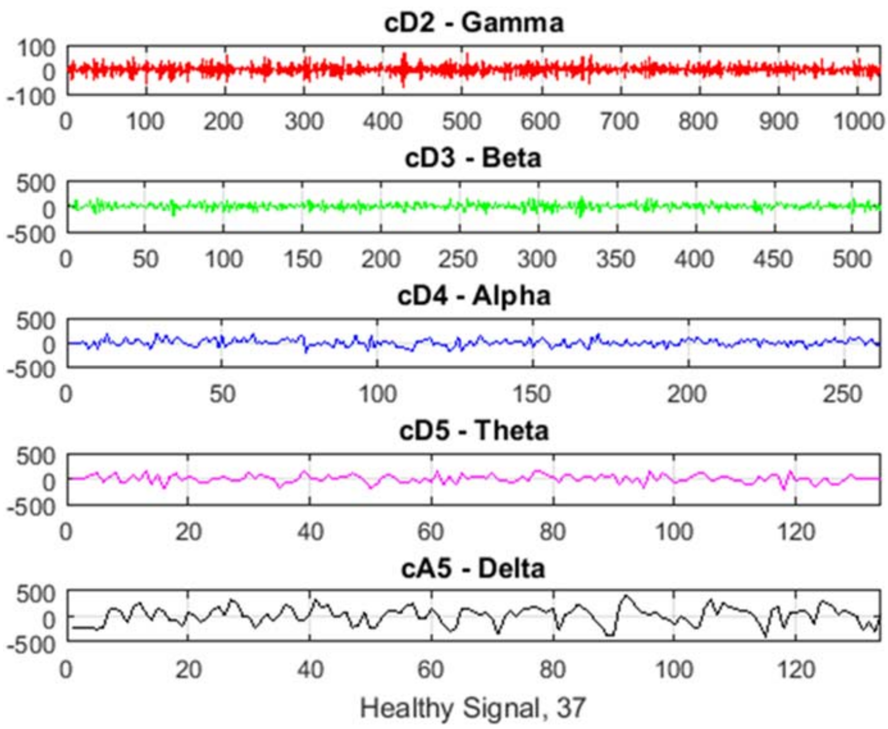

Figure 4.5: Brain rhythm analysis using db4 (window 37).

Then 4097 coefficients of the wavelet were taken for each EEG window. Therefore, the time domain of the signal should be transformed to the frequency domain in order to obtain further information about the signal features. The wavelet transform is able to acquire the information in the time or frequency domain. Thus, it gives the user a choice of the appropriate information they require. Moreover, it is useful in eliminating any Electromyography (EMG) and Electrooculography (EOG) artifacts in the EEG signals.

In this study, the following statistical parameters were applied to describe the time-frequency distribution of the EEGs to minimize the dimensionality of the extracted vectors of the feature. The statistics on the set of wavelet coefficients were utilized from each sub band and selected to know the type EEG signal as follow:

1) Maximum of the wavelet coefficients in each sub band.
2) Minimum of the wavelet coefficients in each sub band.
3) Mean of the wavelet coefficients in each sub band is obtain from the equation:

$$\mu_i = \frac{1}{n}\sum_{j=1}^{n} Dij \qquad i = 1,2,3 \ldots\ldots\ldots..I \tag{4.2}$$

4) The standard deviation (SD) of the wavelet coefficients in each sub band:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{I=1}^{N}(D_i - \mu)^2} \tag{4.3}$$

5) Entropy in the sub band is a mathematical pattern of the distrust of the result wherever the signal included a thousand folds of bits of data. The analytical description is

$$Entropy\ (EN) = \sum_{j=1}^{N} Di^2\, jlog(Di^2 j) \qquad i = 1,2,3 \ldots\ldots..I \tag{4.4}$$

6) The energy of the wavelet coefficients in the each sub band shows the power of the signal since it provides the region below the curve line of power in any period of time. The energy of the EEGs of the limited period is given as:

$$Energy\ (EN) = \sum_{i=1}^{N} |D_{ij}|^2 \qquad i = 1,2,3 \ldots\ldots..I \tag{4.5}$$

7) The medium power in each sub band for the wavelet coefficients.
8) The rate of the absolute mean values nearby in each sub band.

The classification results will be shown in the output of ANN. Figure 4.6. shows the schematic of the proposed neural network model in case 1 as it will be explained later in this section.

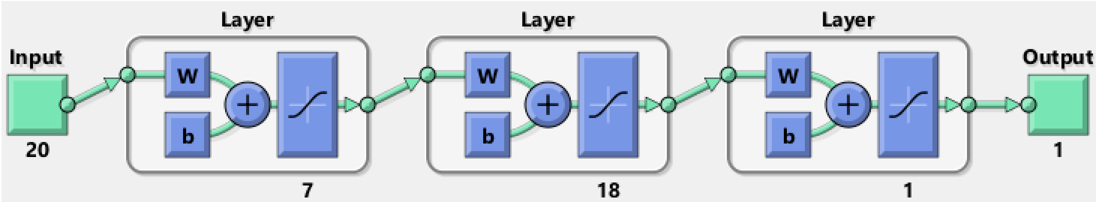

Figure 4.6: Schematic of the proposed neural network model.

The final step will be the evaluating step. Here the accuracy of the healthy (Specificity) and non-healthy (Sensitivity) EEG patterns will be calculated using Confusion Matrix equations as shown in Figure 4.7.

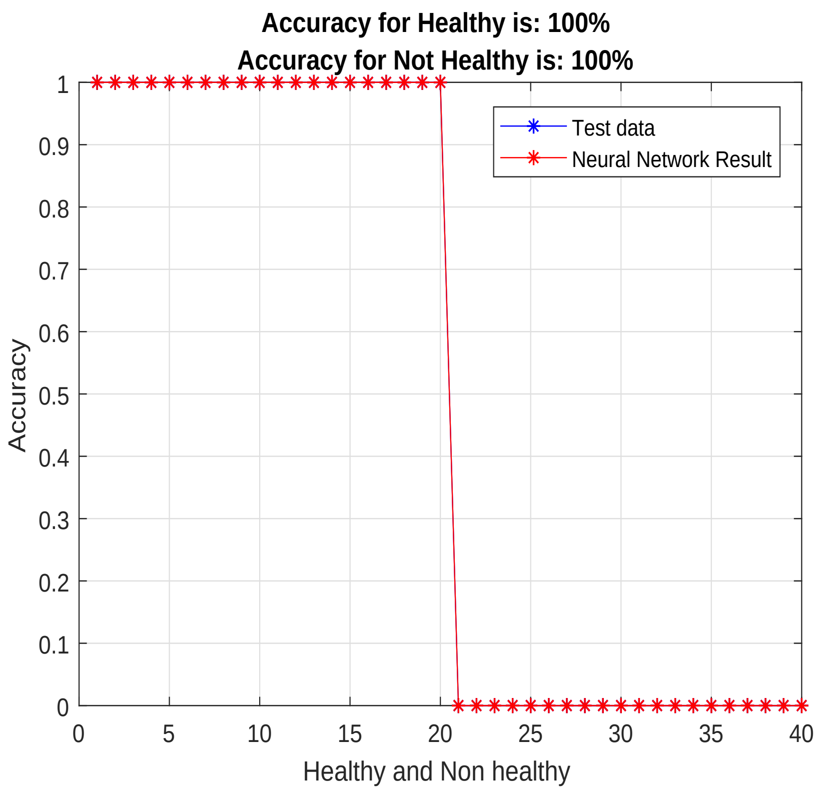

Figure 4.7: Accuracy of the healthy and non-healthy.

4.5 Results

In this stage, the EEG signal is decomposed into five sub-bands: Delta, Theta, Alpha, Beta, and Gamma according to the frequencies mentioned in Table 4.2. Then, we extract the 8 features of each sub-bands in as shown in Table 4.3. Here, the total extracted features from five levels of the analyzed windows is equal to

5 levels * 8 features * 200 windows = 8000 features

Table 4.3: The extracted Features in the decomposed levels.

#	Gamma	Beta	Alpha	Theta	Delta
1	Mean values	Mean values	Mean values	Mean values	Mean values
2	STD	STD	STD	STD	STD
3	Band Power	Band Power	Band Power	Band Power	Band Power
4	Absolute Mean Values	Absolute Mean Values	Absolute Mean Values	Absolute Mean Values	Absolute Mean Values
5	Max	Max	Max	Max	Max
6	Min	Min	Min	Min	Min
7	Entropy	Entropy	Entropy	Entropy	Entropy
8	Energy	Energy	Energy	Energy	Energy

These features are estimated via the Kruskal-Wallis statistical method which can be applied to decide whether there are statistically important differences between two or more groups of an independent variable.

In this step, we discuss six different cases of the extracted features that will be delivered to classification stage in multi scenarios. In the first case, we consider the most 4 features used in the previous studies i.e. mean, rate power, standard deviation, and rate of the absolute mean values. These features will be referred as "essential features". The remaining 4 features, called "secondary features", will be considered in the others cases by combining them with the essential features as shown in Table 4.4. In this table, the (✓) sign indicates the using of feature in classification. However, the (*) sign indicates that the corresponding feature is not used in classification.

Table 4.4: Different cases of classification.

Feature / Case	Mean	Avg. Power	deviation.	Ratio Mean	Max.	Min.	Entropy	Energy
Cas.1	✓	✓	✓	✓	*	*	*	*
Cas.2	✓	✓	✓	✓	✓	*	*	*
Cas.3	✓	✓	✓	✓	*	✓	*	*
Cas.4	✓	✓	✓	✓	*	*	✓	*
Cas.5	✓	✓	✓	✓	*	*	*	✓
Cas.6	✓	✓	✓	✓	✓	✓	✓	✓

The resulted vector of all previous cases is implemented to Multilayer Perceptron Neural Network (MLPNN) with an Levenberg-Marquardt (LM) algorithm for the classification.

80% of dataset is employed for running (training neurons) and the 20% of dataset is used for validation (testing), as shown in Table 4.5.

Table 4.5: Distribution of datasets A & Dataset D.

Type	Training Dataset	Test Dataset	Complete Datasets
Non-Healthy	80	20	100
Healthy	80	20	100
Complete Datasets	160	40	200

The output decisions of ANN stage are evaluated by using the Confusion Matrix equations in order to estimate the Specificity, the Sensitivity and the Accuracy of our algorithm as shown in the following paragraphs.

In the first case, using the 4 essential features, we have repeated the classification process ten times before calculating the average specificity and the average sensitivity. This will allow us to find accuracy ratio of our algorithm using both function db4 and db8 as shown in Figure 4.8 and Figure 4.9.

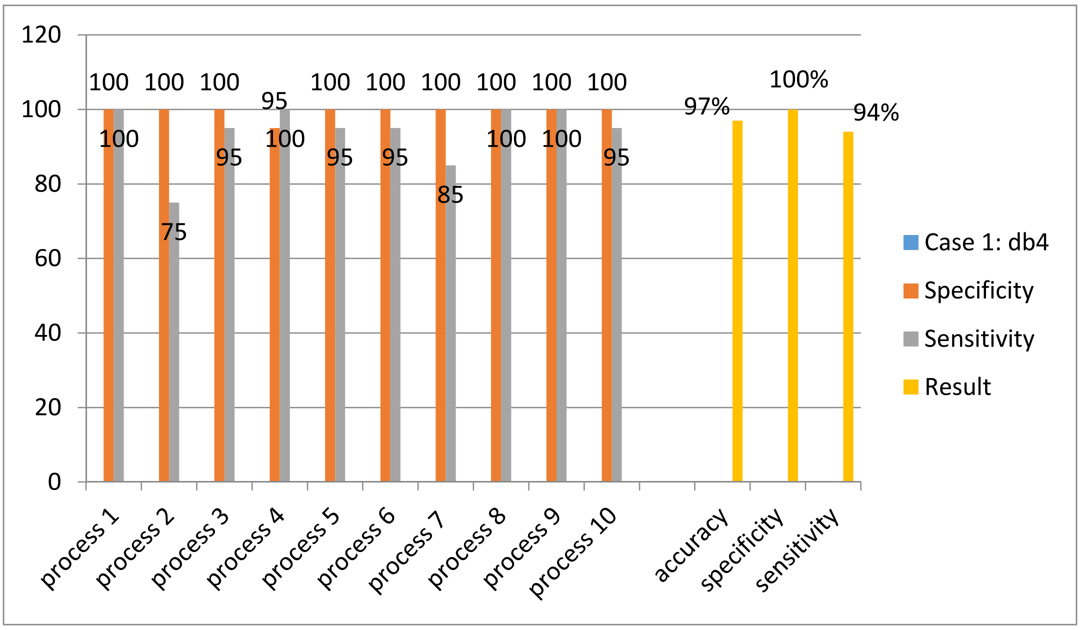

Figure 4.8: Average accuracy of case 1 using db4 (Essential features).

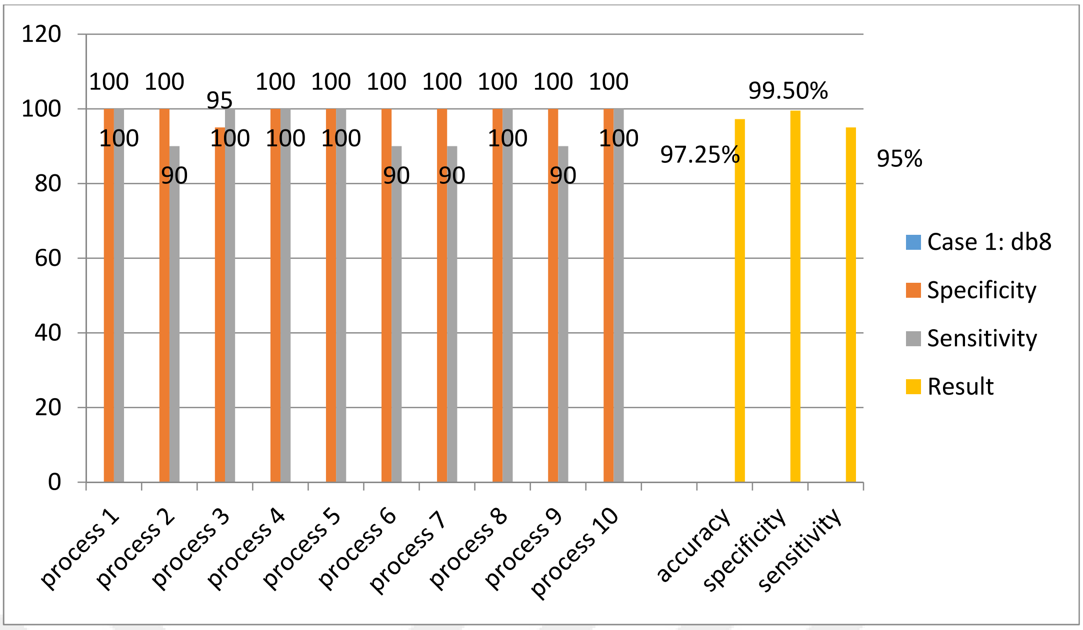

Figure 4.9: Average accuracy of case 1 using db8 (Essential features).

The same procedure was repeated to measure the accuracy of our proposed algorithm in other cases.

Figure 4.10 and Figure 4.11 show the evaluation of the second case using the 4 essential features plus the maximum feature using both function db4 and db8.

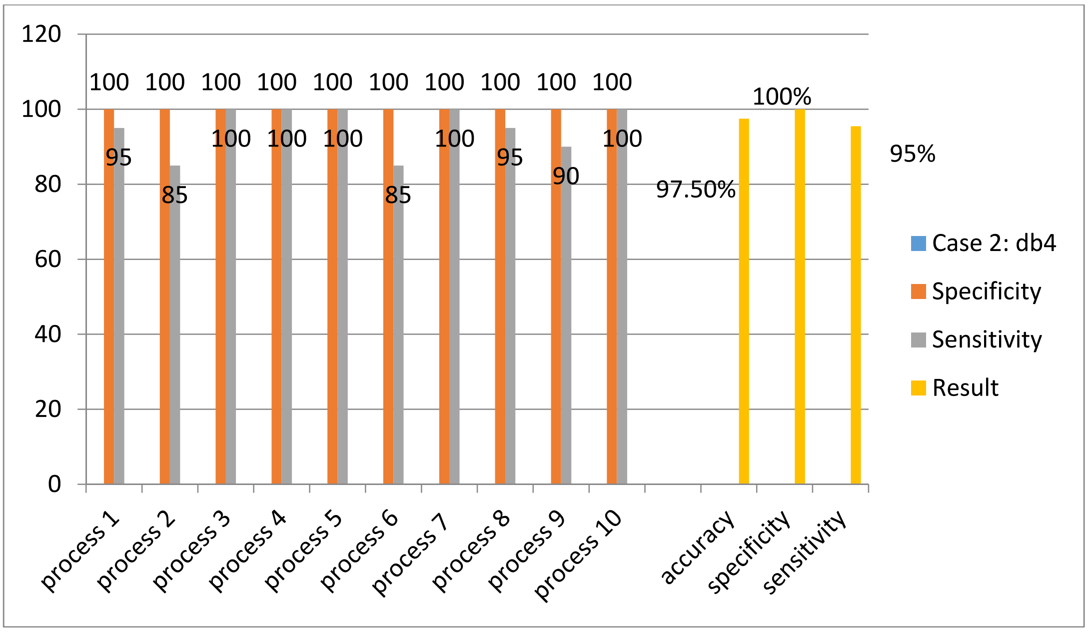

Figure 4.10: Average accuracy of case 2 using db4.

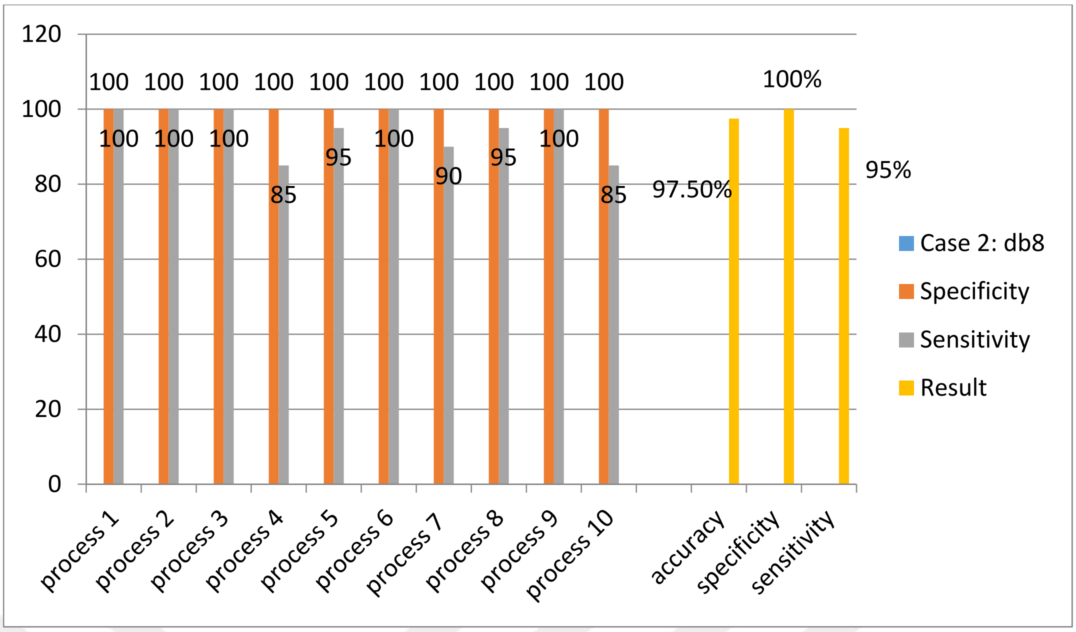

Figure 4.11: Average accuracy of case 2 using db8.

Figure 4.12 and Figure 4.13 show the evaluation of the third case using the 4 essential features plus the minimum feature using both function db4 and db8.

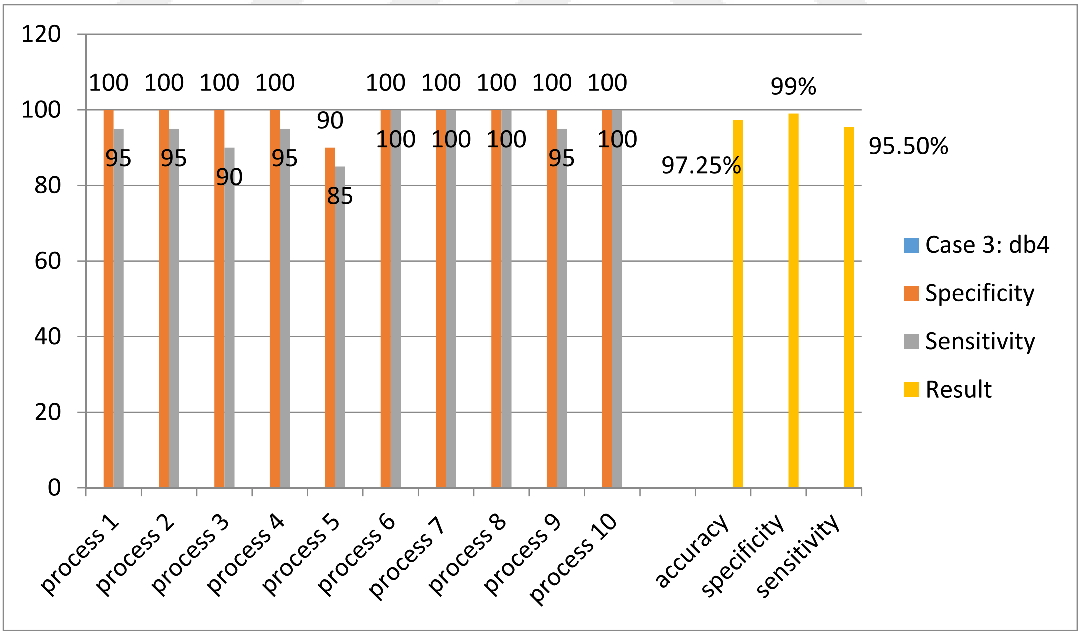

Figure 4.12: Average accuracy of case 3 using db4.

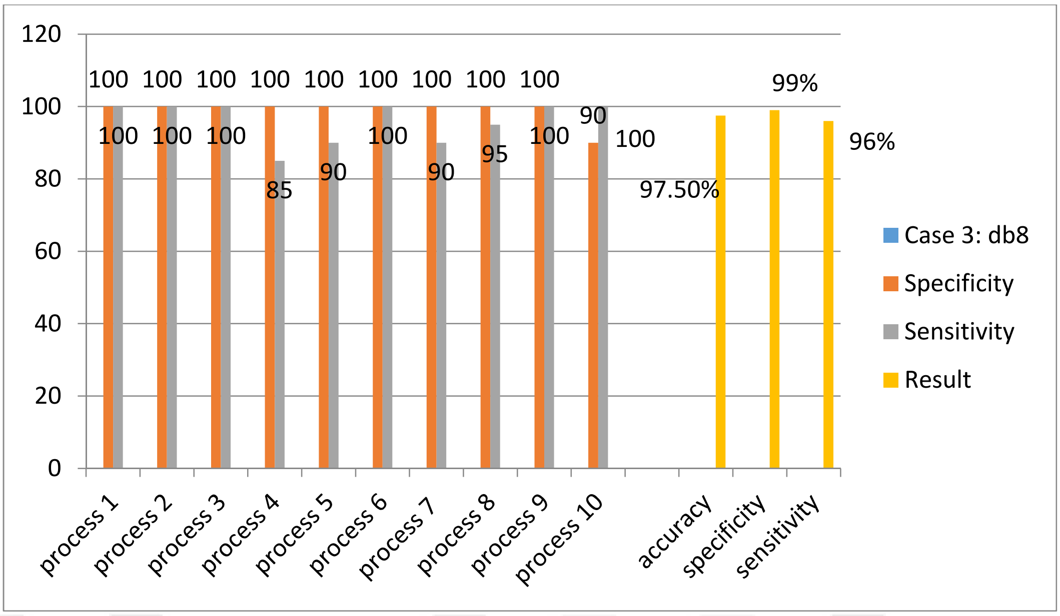

Figure 4.13: Average accuracy of case 3 using db8.

Figure 4.14 and Figure 4.15 show the evaluation of the **fourth case** using the 4 essential features plus the entropy feature using both function db4 and db8.

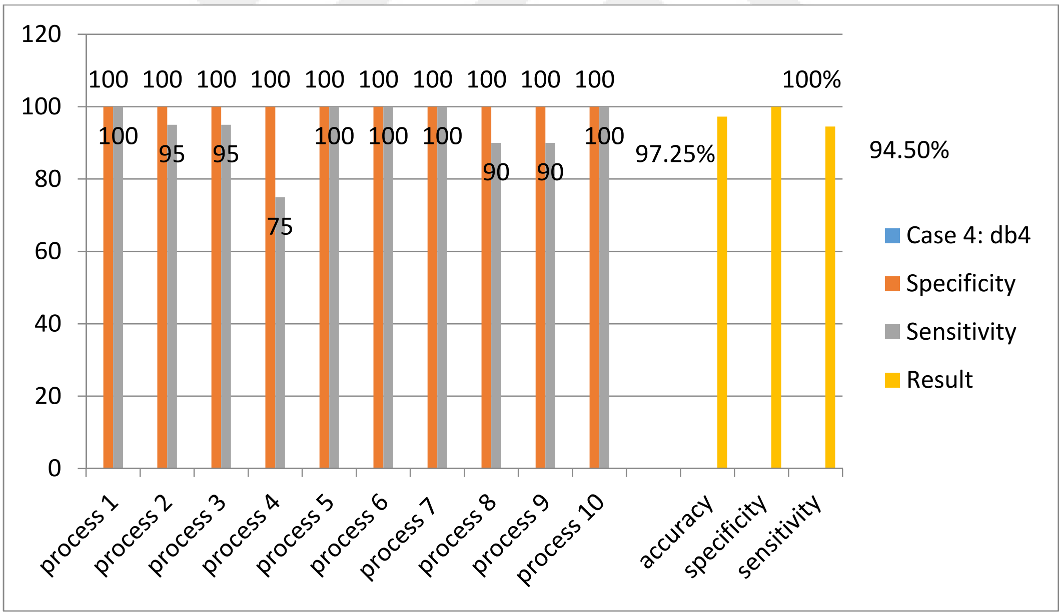

Figure 4.14: Average accuracy of case 4 using db4.

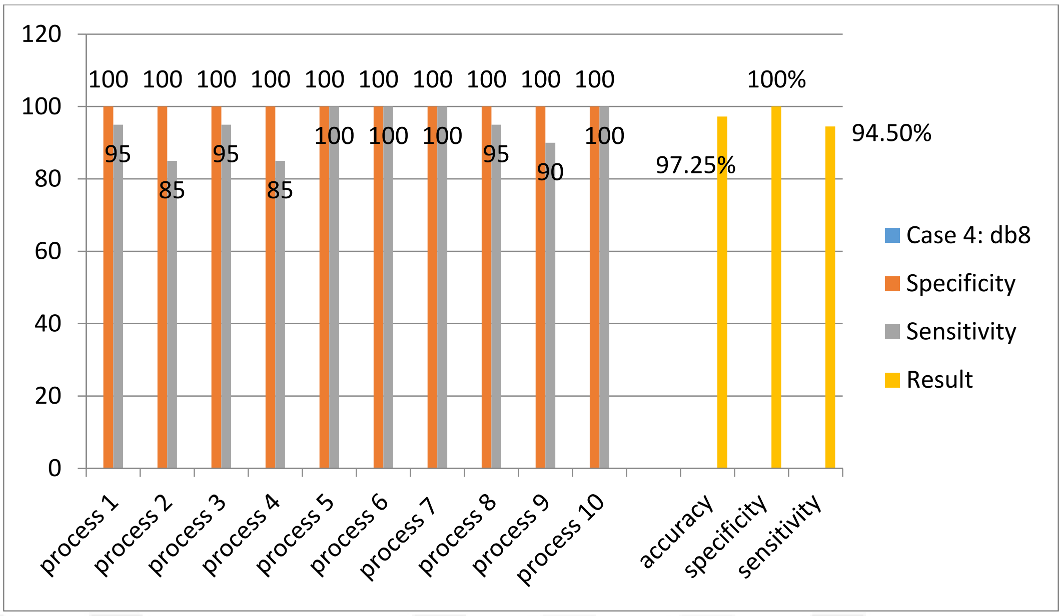

Figure 4.15: Average accuracy of case 4 using db8.

Figure 4.16 and Figure 4.17 show the evaluation of the fifth case using the 4 essential features plus the energy feature using both function db4 and db8.

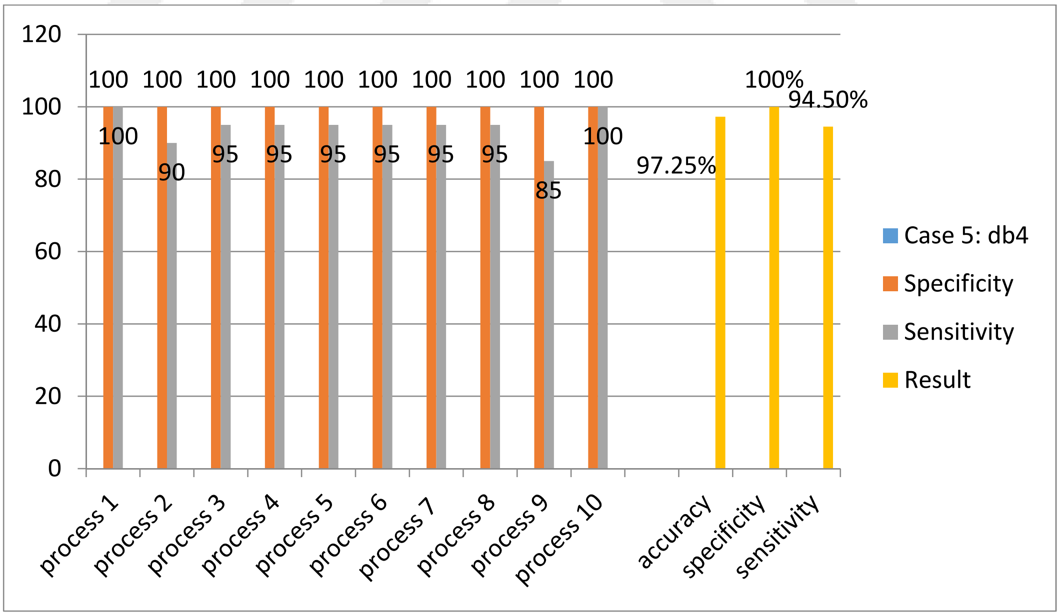

Figure 4.16: Average accuracy of case 5 using db4.

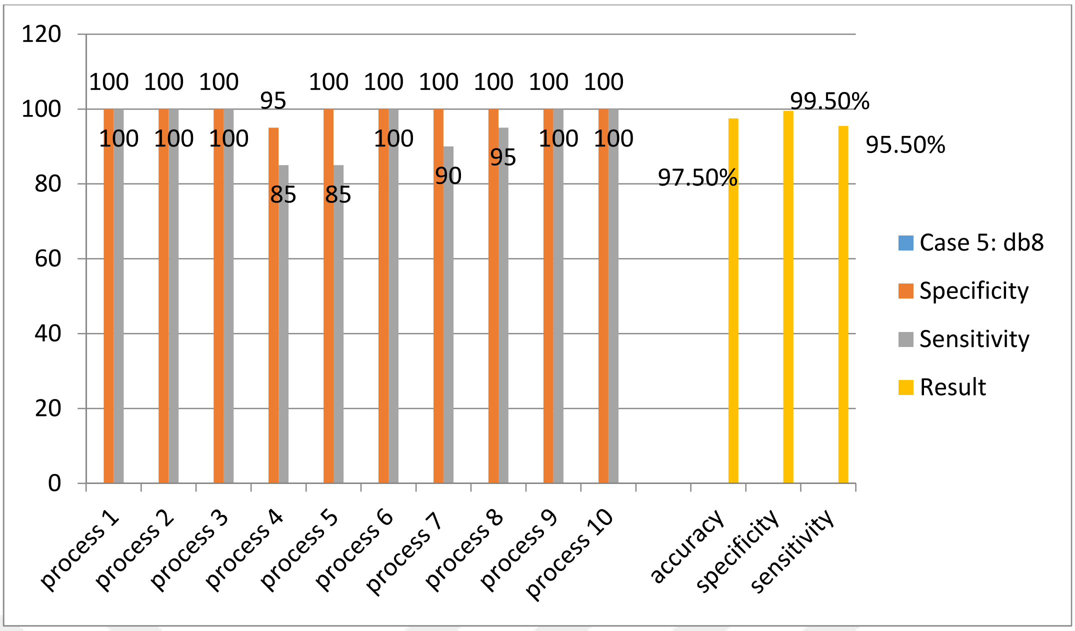

Figure 4.17: Average accuracy of case 5 using db8.

The previous results show the impact of each secondary feature individually. However, Figure 4.18 and 4.19_shows the last scenario which correspond to the accuracy result of all features (i.e secondary plus essential features).

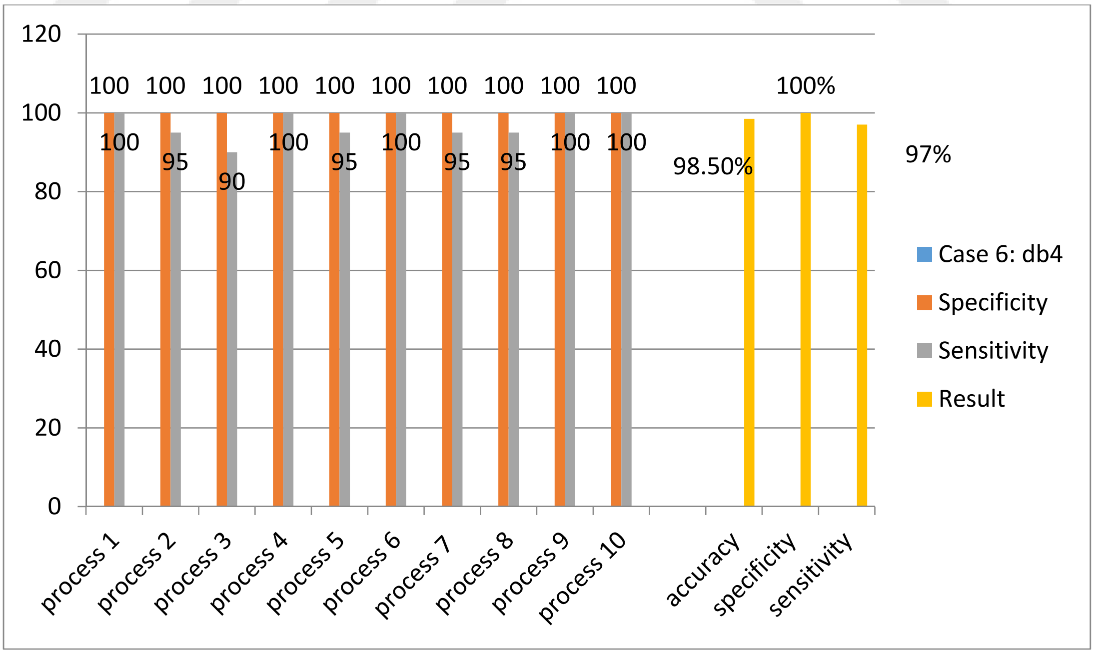

Figure 4.18: Average accuracy of case 6 using db4.

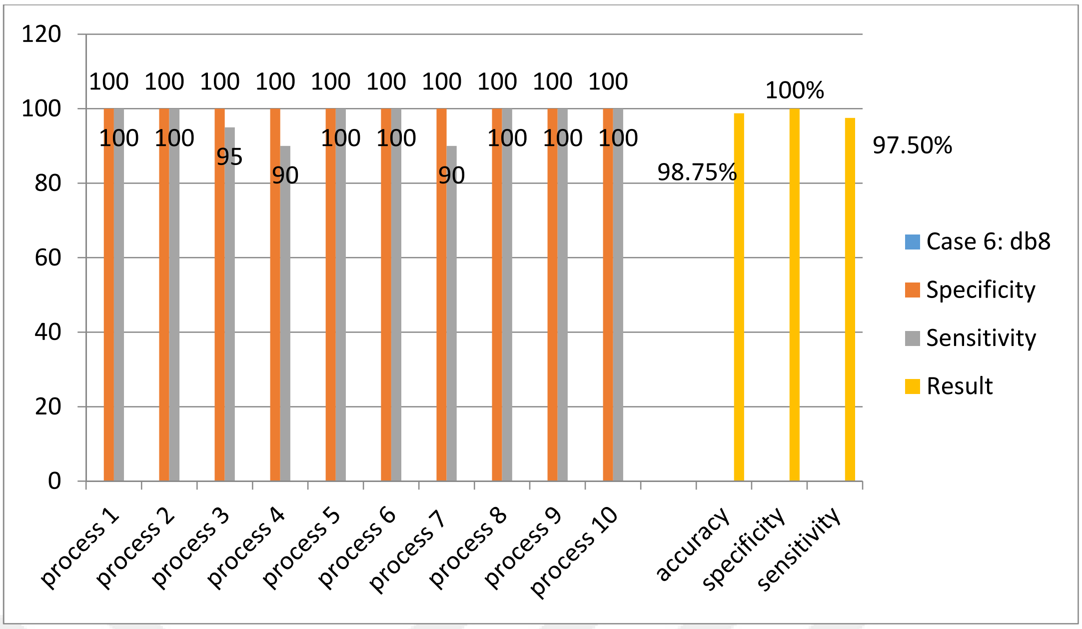

Figure 4.19: Average accuracy of case 6 using db8.

Table 4.6 summarizes the evaluation results of our algorithm for all discussed scenarios. This table also show the accuracy increasing percentage resulting from adding each secondary feature to the essential one (case 1).

Table 4.6: Comparison between the six cases using db4 and db8.

Cases	Accuracy %	Sensitivity %	Specificity %	Wavelet function	Percent Increase
Case 1	97%	94%	100%	db4	---
Case 2	97.50%	95%	100%		0.50%
Case 3	97.25%	95.5%	99%		0.25%
Case 4	97.25%	94.5%	100%		0.25%
Case 5	97.25%	94.5%	100%		0.25%
Case 6	98.50%	97%	100%		1.50%
Case 1	97.25%	95%	99.50%	db8	---
Case 2	97.50%	95%	100%		0.25%
Case 3	97.50%	96%	99%		0.25%
Case 4	97.25%	94.50%	100%		0%
Case 5	97.50%	95.50%	99.50%		0.25%
Case 6	98.75%	97.50%	100%		1.50%

Table 4.7 and Figure 4.20_show the comparison between our proposed method and previous works that used the same database. It is clear that our proposed algorithm outperform the pervious developped methods.

Table 4.7: Comparing previous studies conducted on the UBonn dataset.

Author	Wavelet Function	Classification	Accuracy	Year
Nunes [10]	DWT (db4)	ANNs	93%	2014
Fathima [11]	DWT (db4)	ANNs	95.3%	2013
Omerhodzic [12]	DWT (db4)	ANNs	94 %	2013
Subasi [14]	DWT (db4)	ANNs	93.2%	2007
Our work	**DWT (db4)**	**ANNs**	**98.50%**	**2017**
Our work	**DWT (db8)**	**ANNs**	**98.75%**	**2017**

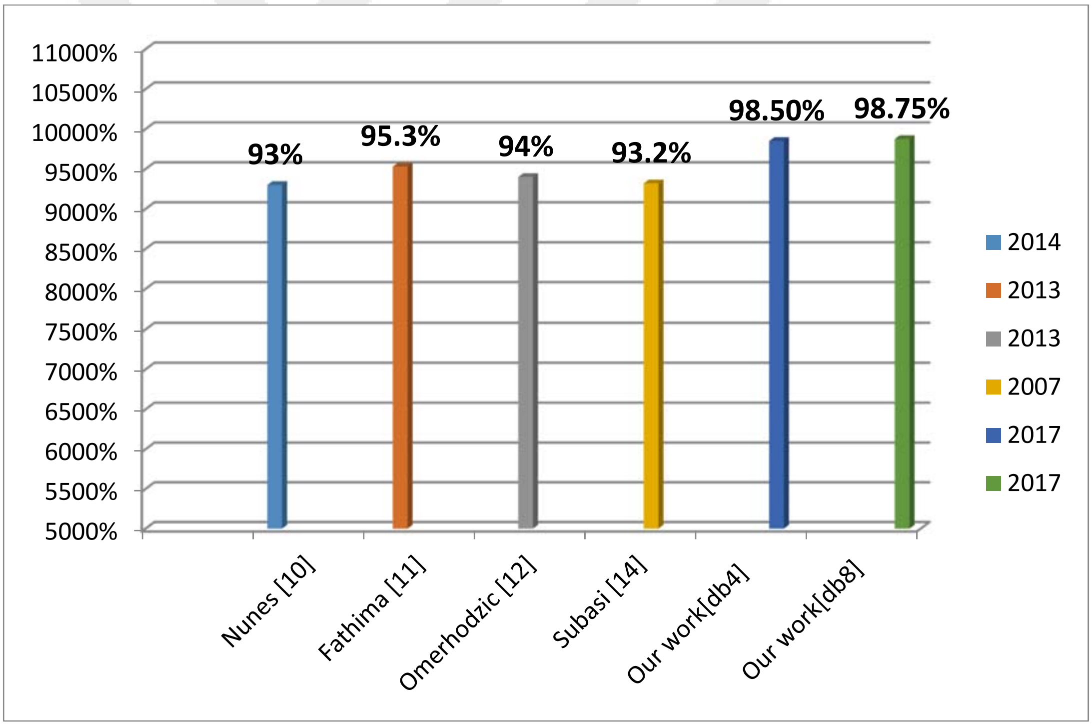

Figure 4.20: Comparison between previous works and this study.

CHAPTER FIVE

CONCLUSIONS & FUTURE WORKS

5.1 Conclusions

In this work, high accurate algorithm is proposed to identify epilepsy seizures in EEG signals. Our study is performed depending on Bonn University Dataset and achieved via MATLAB environment using MATLAB Wavelet Toolbox. In the pre-processing phase, the EEG signal is decomposed into five levels by means of Discrete Wavelet Transform using 2 functions, db4 and db8.

In order to study the impact of the secondary features on the classification accuracy, six different classification cases have been discussed. The resulted features vector in all scenarios was input to 2 hidden layers of ANN trained by Levenberg-Marquardt method to deliver the classification decision.

The proposed method was evaluated by Confusion Matrix equations. It was clear that our algorithm outperforms the previous studies applied on the same database.

The best classification accuracy obtained when extracting eight features from the EEG signals was 98.50% in Daubechies order 4 and 98.75% in Daubechies order 8 respectively.

5.2 Future Works

The perspective of this work can be using the Field Programmable Gate Array (FPGA) to implement the proposed method to find healthy and non-healthy EEG signals. For feature extraction, we can use the entropy of the signals at different frequencies.

REFERENCES

[1] D. Freeman, C. J. Bajema, J. Blacking, R. L. Carneiro, U. Cowgill, S. Genovés*, et al.*, "The Evolutionary Theories of Charles Darwin and Herbert Spencer," *Current Anthropology,* vol. 15, pp. 211-237, 1974.

[2] S. Finger, *Origins of neuroscience: a history of explorations into brain function*: Oxford University Press, USA, 2001.

[3] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, *Principles of neural science* vol. 4: McGraw-hill New York, 2000.

[4] D. Chen, S. Wan, and F. S. Bao, "Epileptic Focus Localization Using Discrete Wavelet Transform Based on Interictal Intracranial EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 25, pp. 413-425, 2017.

[5] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 24, pp. 28-35, 2016.

[6] G. Singh, M. Kaur, and D. Singh, "Detection of an epileptic seizure using wavelet transformation and spike based features," in *Recent Advances in Engineering & Computational Sciences (RAECS), 2015 2nd International Conference on*, 2015, pp. 1-4.

[7] K. Abualsaud, M. Mahmuddin, M. Saleh, and A. Mohamed, "Ensemble classifier for epileptic seizure detection for imperfect EEG data," *The Scientific World Journal,* vol. 2015, 2015.

[8] Y. Kumar, M. Dewal, and R. Anand, "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network," *Signal, Image and Video Processing,* vol. 8, pp. 1323-1334, 2014.

[9] B. S. Nanthini and B. Santhi, "Different approaches to analyzing EEG signals for seizure detection," *International Journal of Signal and Imaging Systems Engineering,* vol. 8, pp. 28-38, 2015.

[10] T. M. Nunes, A. L. Coelho, C. A. Lima, J. P. Papa, and V. H. C. de Albuquerque, "EEG signal classification for epilepsy diagnosis via optimum path forest–A systematic assessment," *Neurocomputing,* vol. 136, pp. 103-123, 2014.

[11] T. Fathima, M. Bediuzzaman, and P. K. Joseph, "Wavelet based features for classification of normal, ictal and interictal EEG signals," *Journal of Medical Imaging and Health Informatics,* vol. 3, pp. 301-305, 2013.

[12] I. Omerhodzic, S. Avdakovic, A. Nuhanovic, and K. Dizdarevic, "Energy distribution of EEG signals: EEG signal wavelet-neural network classifier," *arXiv preprint arXiv:1307.7897,* 2013.

[13] E. D. Übeyli, "Combined neural network model employing wavelet coefficients for EEG signals classification," *Digital Signal Processing,* vol. 19, pp. 297-308, 2009.

[14] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of the expert model," *Expert Systems with Applications,* vol. 32, pp. 1084-1093, 2007.

[15] H. R. Mohseni, A. Maghsoudi, M. H. Kadbi, J. Hashemi, and A. Ashourvan, "Automatic Detection of Epileptic Seizure using Time-Frequency Distributions," in *Advances in Medical, Signal and Information Processing, 2006. MEDSIP 2006. IET 3rd International Conference On*, 2006, pp. 1-4.

[16] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *IEEE Transactions on Biomedical Engineering,* vol. 54, pp. 205-211, 2007.

[17] A. C. Merzagora, S. Bunce, M. Izzetoglu, and B. Onaral, "Wavelet analysis for EEG feature extraction in deception detection," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, 2006, pp. 2434-2437.

[18] M. Kalaivani, V. Kalaivani, and V. A. Devi, "Analysis of EEG Signal for the Detection of Brain Abnormalities," *at International Journal of Computer Applications® year,* 2014.

[19] T. Ahmed, M. Islam, M. S. U. Yusuf, and M. Ahmad, "Wavelet based analysis of EEG signal for evaluating mental behavior," in *Informatics, Electronics & Vision (ICIEV), 2013 International Conference on*, 2013, pp. 1-6.

[20] H. Sharabaty, "Diagnostic de la somnolence d'un opérateur: analyse automatique de signaux physiologiques," Phd thesis, Université Paul Sabatier - Toulouse III, France, 160p, 2007.

[21] H. Sharabaty, B. Jammes, D. Esteve, "EEG analysis using HHT: One step toward automatic drowsiness scoring", in Proc. 22nd Int. Conf. on Advanced Information Networking and Applications, Okinawa, Japan, 2008 pp. 826–831.

[22] B. Jammes, H. Sharabaty, and D. Esteve, "Alpha and theta wave localisation in EEG signals: survey of existing algorithms", LAAS-CNRS, Report 05467, European Project IST SENSATION N°507231 17p, Toulouse, 2005.

[23] G. Strang and T. Nguyen, *Wavelets and filter banks*: SIAM, 1996.

[24] J.-S. Guf and W.-S. Jiang, "The Haar wavelets operational matrix of integration," *International Journal of Systems Science,* vol. 27, pp. 623-628, 1996.

[25] Y. Meyer, *Wavelets and operators* vol. 1: Cambridge university press, 1995.

[26] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on signal processing,* vol. 41, pp. 3397-3415, 1993.

[27] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on pure and applied mathematics,* vol. 41, pp. 909-996, 1988.

[28] A. Mertins, "Signal Analysis: Wavelets, Filter Banks, Time-Frequency Transforms and Applications, English (revised edition)," 1999.

[29] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM journal on mathematical analysis,* vol. 15, pp. 723-736, 1984.

[30] P. Addison, J. Watson, and T. FENG, "Low-oscillation complex wavelets," *Journal of Sound and Vibration,* vol. 254, pp. 733-762, 2002.

[31] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence,* vol. 11, pp. 674-693, 1989.

[32] R. Ganesan, T. K. Das, and V. Venkataraman, "Wavelet-based multiscale statistical process monitoring: A literature review," *IIE transactions,* vol. 36, pp. 787-806, 2004.

[33] D. E. Dutkay and P. E. Jorgensen, "Wavelets on fractals," *Revista Matemática Iberoamericana,* vol. 22, pp. 131-180, 2006.

[34] J. Rafiee, P. Tse, A. Harifi, and M. Sadeghi, "A novel technique for selecting mother wavelet function using an intelli gent fault diagnosis system," *Expert Systems with Applications,* vol. 36, pp. 4862-4875, 2009.

[35] S. Haykin, *Neural networks: a comprehensive foundation*: Prentice Hall PTR, 1994.

[36] S. Haykin and N. Network, "A comprehensive foundation," *Neural Networks,* vol. 2, p. 41, 2004.

[37] D. Saad, *On-line learning in neural networks* vol. 17: Cambridge University Press, 2009.

[38] https://medium.com/autonomous-agents/mathematical-foundation-for-activation-functions-in-artificial-neural-networks-a51c9dd7c089.

[39] http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3.

[40] Y.-L. Wu, D. Agrawal, and A. El Abbadi, "A comparison of DFT and DWT based similarity search in time-series databases," in *Proceedings of the ninth international conference on Information and knowledge management*, 2000, pp. 488-495.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name	: Omar Saadi Fathi ALSHEAR
Nationality	: IRAQI
Date and Place of Birth	: IRAQ- Ninawa, 1982-5-30
Phone	: 00905365810522
Email	: omarsaad529@gmail.com

EDUCATION

Degree	Institution	Year of Graduation
BS	University of Mosul- College of Engineering	2006
High School	Alrisala	2001

WORK EXPERIENCE

Year	Place	Enrollment
2007 - 2015	Iraqi Ministry of Electricity	Maintenance Engineer

FOREIGN LANGUAGES

English, Turkish