Chapter 3

Proposed Model for Analysing EEG Signals

This chapter presents the description of proposed model for classification of three classes of EEG signals: Ictal, Interictal and Healthy. We explain the methodology used to analyze and classify EEG signals related to stages of epilepsy. It also presents the characteristics of EEG database used in this research.

3.1 Introduction

The general block diagram of the proposed approach to analyze and classify EEG signals is shown in Figure 3.1, which is divided into three modules: preprocessing, feature extraction and multi-class classification. Different filters were used for preprocessing the EEG database, whereas Wavelet Transforms (WT) were used to extract features of EEG signal. In this work, delta (δ) and alpha (α) sub-bands were used to characterize the EEG, because some researches have reported that these sub-bands provide useful information to localize a seizure [SUN12], [RAV12]. The proposed model, called MRW-FFWNN, was used to classify epileptic seizure using two classification strategies, called binary-tree [MIT97] and OVO [GAL11]. Next,

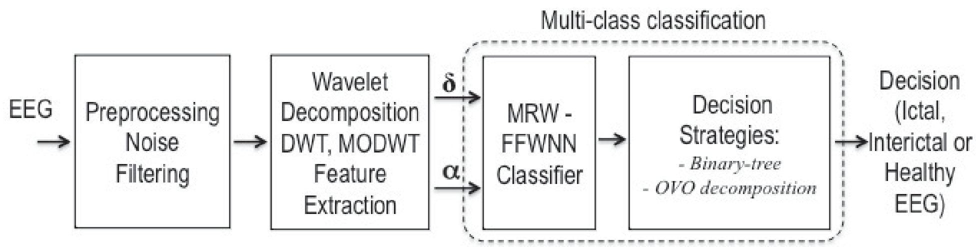

Figure 3.1: General block diagram of proposed approach.

we explain each module of this general block diagram.

3.2 Experimental Data EEG.

As we stated before, the EEG signals used in the experiments reported in this work come from a free-available EEG database, provided by the University of Bonn. The database contains EEG signals from three different events, namely, healthy subjects (healthy EEG), epileptic subjects during seizure-free intervals (interictal EEG) and epileptic subjects during a seizure (ictal EEG) [BON16], [AND01]. This collection contains unfiltered EEGs signals of five subjects, recorded with a sampling rate of 173.6 Hz. The collection contains five datasets identified as: O, Z, F, N and S; each set holds 100 segments of EEG signals of 23.6 seconds, so each segment contains 4,096 samples, with the characteristics described in Table 3.1. Sets O and Z were obtained from healthy subjects with eyes open and closed respectively; sets F and N were obtained during interictal states in different zones of the brain and set S was gotten from an subject during ictal state [TZA09]. In order to make a fair comparison with some of the previous works, sets Z, F and S were used for the results reported in this thesis.

Table 3.1: Summary of the EEG data collection provided by the University of Bonn [BON16], [AND01].

Set	Patient's state	Subject	Electrode type	Electrode-placement	Number samples / Sample duration(s)
Z	Awake and eyes open (Healthy)	Five healthy subjects	Extra-cranial	International 10-20 system	100 / 23.6
O	Awake and eyes closed (Healthy)	Five healthy subjects	Extra-cranial	International 10-20 system	100 / 23.6
N	Seizure-free (Inter-ictal)	Five epileptic patients	Intra-cranial	Opposite to epileptogenic zone	100 / 23.6
F	Seizure-free (Inter-ictal)	Five epileptic patients	Intra-cranial	Within epileptogenic zone	100 / 23.6
S	Seizure activity (Ictal)	Five epileptic patients	Intra-cranial	Within epileptogenic zone	100 / 23.6

3.3 Preprocessing.

The EEG signals were preprocessed as reported in our previous publications [JUA13a], [GOM14], [JUA13b]. A filtering of the EEG signals was performed in order to remove noise added during recording. Some physiological researchers consider that EEG frequencies above 60 Hz are noise and can be neglected [MIR11]. Considering this value, the cut-off frequency of the low-pass filters used here was set to 64 Hz. The value 64, which is an exact power of two, was used instead of 60 Hz, in order to simplify the calculation of the frequency sub-bands of the EEG during the wavelet analysis. Two approaches for filtering were tested, namely, Finite Impulse Response (FIR) and Infinite Impulse response (IIR). A FIR filter is one whose impulse response (or response to any finite-length input) is of finite duration, because it settles to zero in finite time. An IIR filter may has internal feedback and responds indefinitely, usually decaying [PRO07]. IIR filters (Chebyshev type II and Elliptic) and FIR filters (Equiriple and Least squares) were designed in order of satisfying 3 dB of ripple in the pass-band from 0 to 64 Hz and at least 60 dB of attenuation in

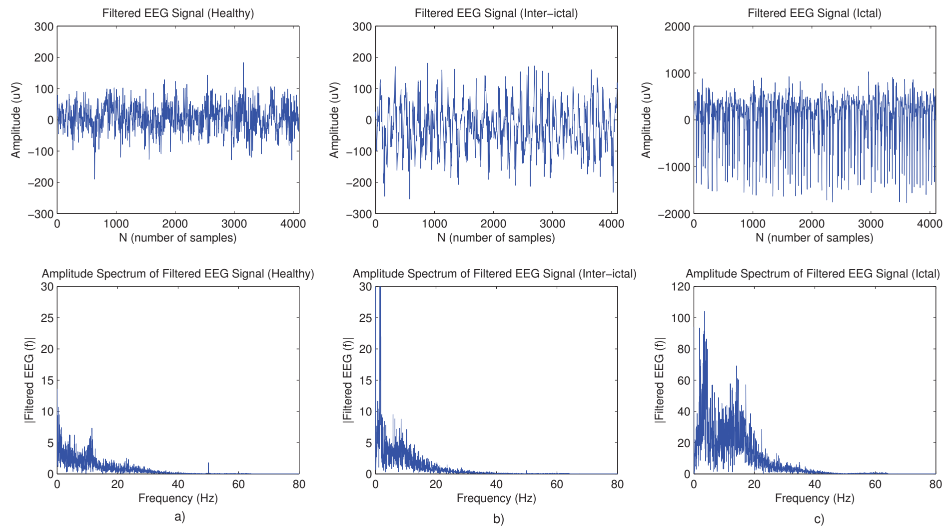

Figure 3.2: Filtered signals EEG by a FIR Least Squares filter and its frequency spectrum of: a) Healthy subject, b) Interictal subject, c) Ictal subject. Upper plots are samples from EEG signals and the lower plots show the frequency components of these EEG signals.

the stop-band using Matlab 2010a and the Signal Processing Toolbox Version 6.19 [PRO07]. The filters obtained with the above characteristics were an IIR Chebyshev type II filter of order 24, an Elliptic filter of order 9, a FIR filter Equiripple of order 343 and a Least Squares filter of order 350 [PRO07].

Figure 3.2 shows a segment of 4,096 samples of a filtered EEG from a healthy subject (part a), an interictal subject (part b) and an ictal subject (part c) with their corresponding frequency spectrum. Notice the differences in the frequency range of each subject. Upper plots of Figure 3.2 show EEG segments while lower plots are the corresponding frequencies. Notice that frequency components above to 64 Hz have been eliminated due to the low-pass filtering.

3.4 Feature extraction.

In this research, the decomposition of EEG signals into delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) sub-bands (see Table 2.1) was carried out using the popular methods of the Discrete Wavelet Transform (DWT) and the Maximal Overlap Discrete Wavelet Transform (MODWT) described in Sections 2.2.4 and 2.2.5. According to the researches Sunhaya, S. et al. [SUN12], and Ravish, D. and Devi S. [RAV12], delta (δ) and alpha (α) sub-bands provide useful information to localize a seizure, due to the rhythmic activity associated with the onset of seizure activity is composed of strong frequency components of these sub-bands. Further details on EEG sub-bands can be found in Section 2.1.3 and [RAV12], [BLI06]. The decomposition and feature extraction of EEG signals was carried out with DWT and MODWT using the wavelets chosen. Finally, each EEG signal was represented by a feature vector of six components, built using the mean, absolute median and variance of both delta and alpha sub-bands.

3.5 Wavelet selection.

One of the most paramount elements that must be considered in wavelet domain studies is the similarity of the signal under investigation with the wavelet to be analyzed. A mother wavelet is said to be similar with a signal, if the wavelet correlation with the signal possess a value near to 1 [SIN06], [KUM13]. Cross correlation is a tool to measure the similarity of two waveforms as a function of a time as it is insensitive to noise [SAR14]. Hence, cross correlation was used to evaluate the degree of similarity between mother wavelets and EEG signals representing ictal, interictal and healthy states, and then the most appropriate mother wavelet was chosen. Using a library provided by Matlab [MAT14], a wavelet filter [MAT14] and each one of the samples of EEG signals from different classes were cross correlated. All the EEG signals of the three class (100 segments of each one) were normalized in $[0, 1]$ before cross correlation. The correlation coefficient may have a value in $[-1, 1]$, where 1 is total positive correlation, 0 is no correlation, and -1 is total negative correlation. A correlation coefficient close to zero indicates a weak linear relationship between two signals, whereas a correlation coefficient of zero would indicate that there is no correlation, or relationship, between two signals. A positive correlation coefficient occurs when the values of both signals increase together, whereas a negative correlation occurs when the increase of one signal corresponds with the decrease of another [MAT14].

We have considered two criteria for the wavelet choice:

- **Criterion 1.** This consists in choosing the wavelet that gives the highest average of the correlation coefficients with each EEG signal.
- **Criterion 2.** This criterion consists in choosing the wavelet that obtained the best correlation coefficient the highest number of times, with each EEG signal.

In this work, we have selected the Daubechies (Db), Coiflet (Coif) and Symlet (Sym) wavelet families for the correlation with the EEG signals. Different orders of Daubechies (Db2, Db4, Db6, Db8 and Db10), Coiflet (Coif1, Coif2, Coif3, Coif4 and Coif5) and Symlet (Sym2, Sym4, Sym6, Sym8 and Sym10) wavelets were used to obtain the correlation coefficients. In this work we used the two criteria for decomposition and feature extraction of EEG signals with DWT and MODWT.

Figures 3.3, 3.4, and 3.5 show the average of the correlation coefficients obtained using the wavelet families filters Daubechies, Coiflet and Symlet, respectively, and the three classes of EEG signals for the wavelet choice using Criterion 1. Figures 3.6, 3.7, and 3.8 show the number of times that the wavelet obtains the best correlation coefficient using the wavelet families filters Daubechies, Coiflet and Symlet, respectively, and the three classes of EEG signals for the wavelet choice by Criterion 2.

According to Figures 3.3, 3.4, and 3.5, considering the best averages of the correlation coefficients for the EEG signals and the wavelet families filter (criterion 1), the feature extraction for the experiments reported here was done using Coiflet 3 (Coif3), Coiflet 5 (Coif5) and Daubechies 6 (Db6) for Ictal, Interictal and Healthy EEG signals, respectively. With respect to the criterion 2, considering Figures 3.6,

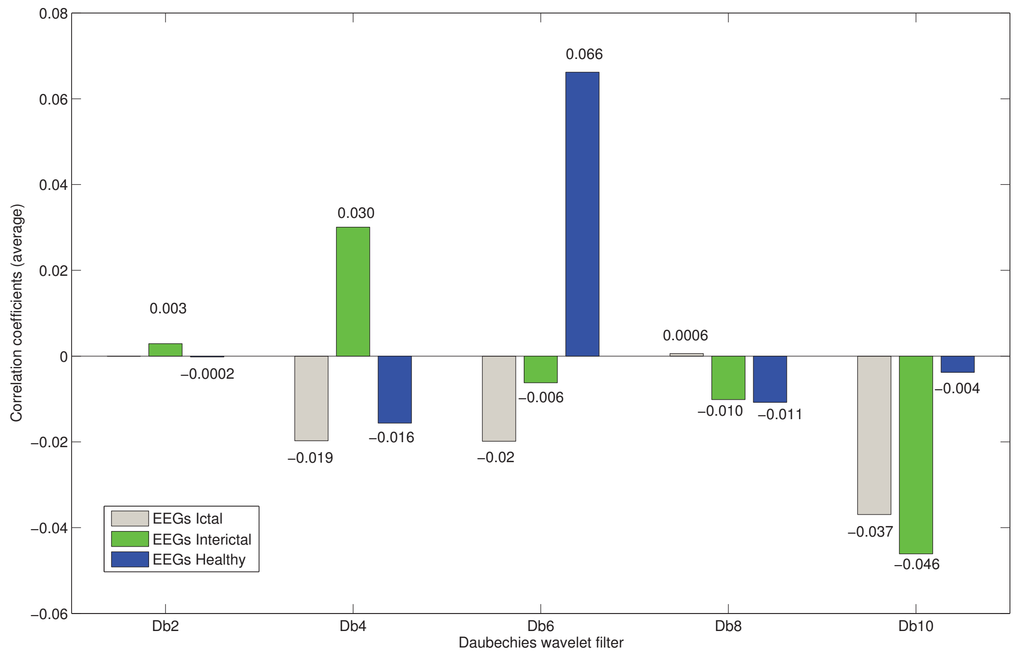

Figure 3.3: Average of correlation coefficients with different orders of Daubechies wavelet filters for Ictal, Interictal and Healthy EEG signals.

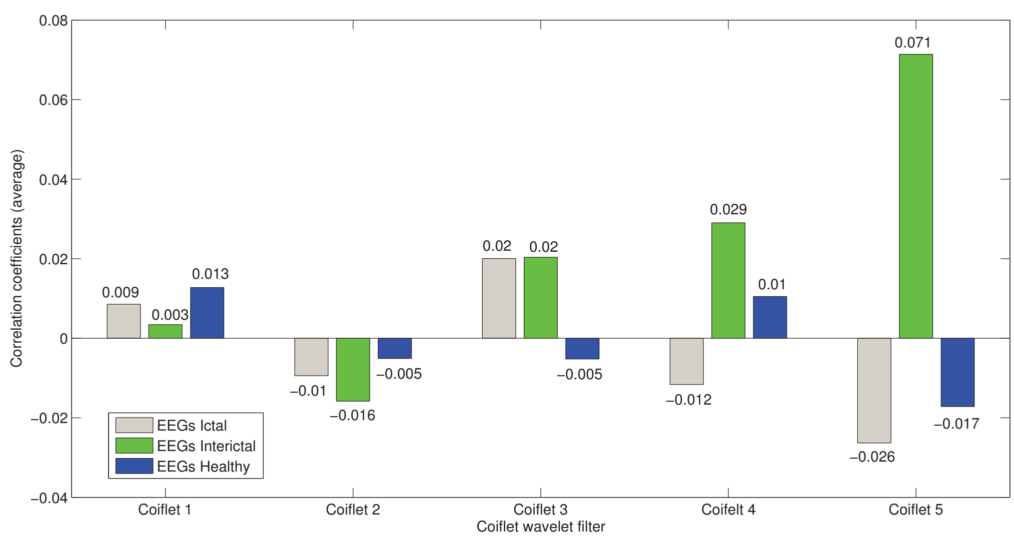

Figure 3.4: Average of correlation coefficients with different orders of Coiflet wavelet filters for Ictal, Interictal and Healthy EEG signals.

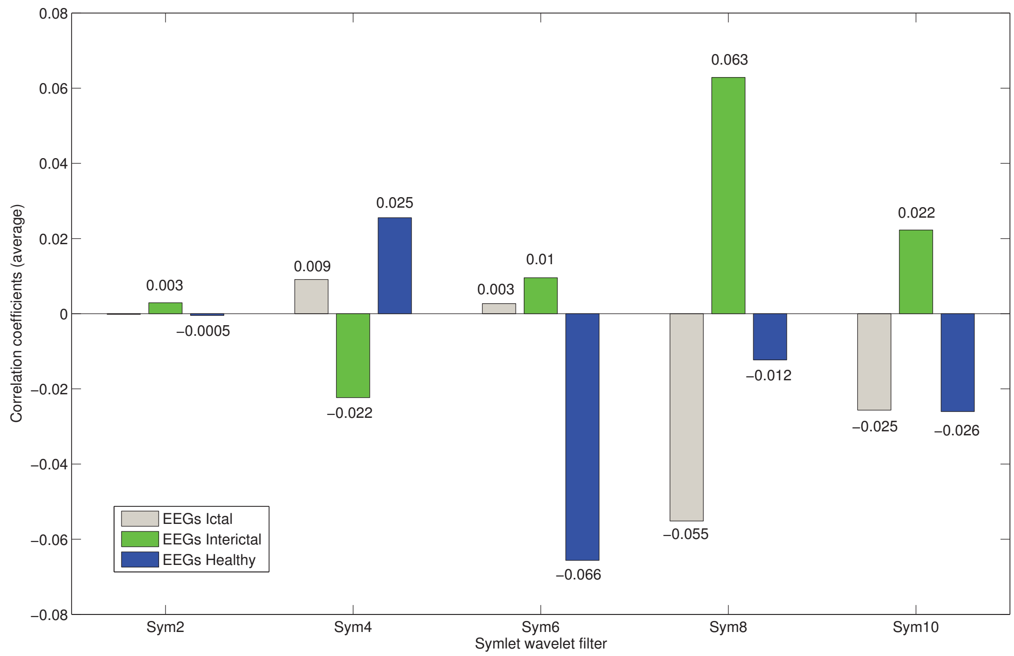

Figure 3.5: Average of correlation coefficients with different orders of Symlet wavelet filters for Ictal, Interictal and Healthy EEG signals.

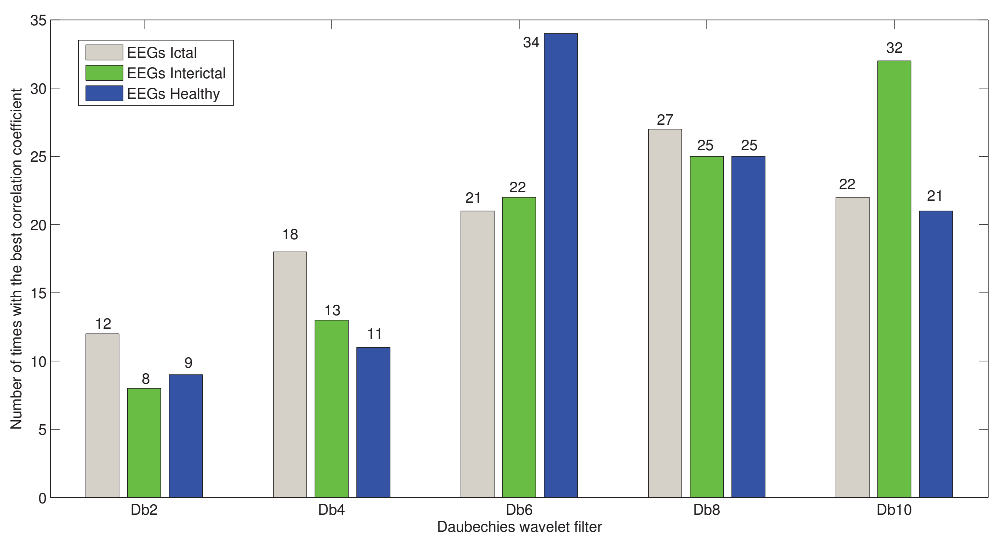

Figure 3.6: Number of times with the best correlation coefficient using different orders of Daubechies wavelet filters and Ictal, Interictal and Healthy EEG signals.

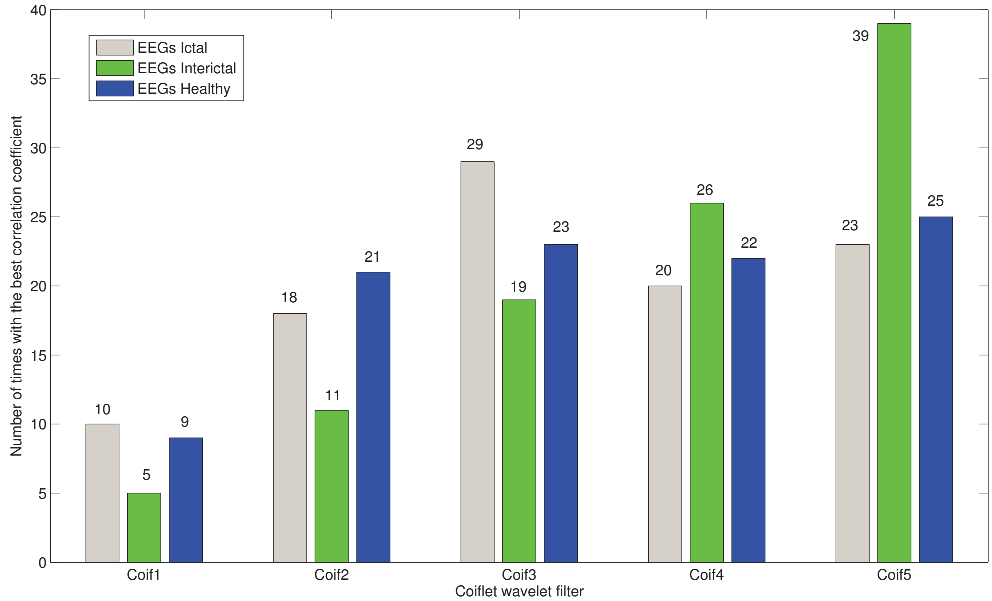

Figure 3.7: Number of times with the best correlation coefficient using different orders of Coiflet wavelet filters and Ictal, Interictal and Healthy EEG signals.

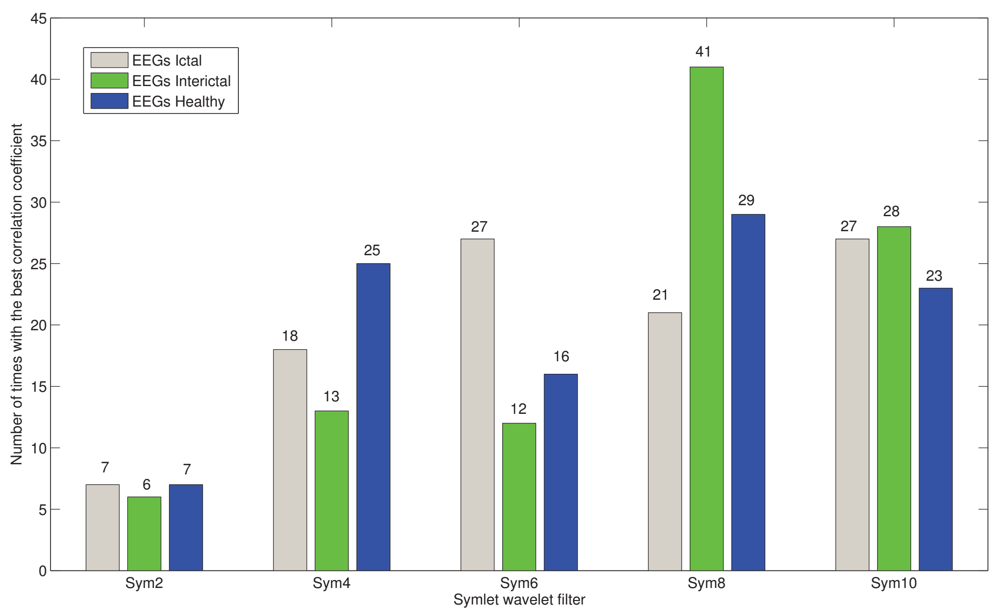

Figure 3.8: Number of times with the best correlation coefficient using different orders of Symlet wavelet filters and Ictal, Interictal and Healthy EEG signals.

Table 3.2: Wavelet selected for feature extraction by criteria 1 and 2.

Criterion	EEG signal	Wavelet chosen	Best parameter
(1) Average of correlation coefficients	Ictal	Coiflet 3 (Coif3)	0.02
	Interictal	Coiflet 5 (Coif5)	0.071
	Healthy	Daubechies 6 (Db6)	0.066
(2) Times number of best correlation coefficient	Ictal	Coiflet 3 (Coif3)	29
	Interictal	Symlet 8 (Sym8)	41
	Healthy	Daubechies 6 (Db6)	34

3.7, and 3.8, and the wavelet that gives the highest number of times with the best correlation coefficient for each one of the EEG signals, the feature extraction was done with Coiflet 3 (Coif3), Symlet 8 (Sym 8) and Daubechies 6 (Db6) for Ictal, Interictal and Healthy EEG signals, respectively. A summary about of wavelet choice for the feature extraction of EEG signals is presented in Table 3.2.

Figures 3.9 and 3.10 show the decomposition of an EEG time series using a four-level DWT and four-level MODWT, respectively, extracting five physiological sub-bands as reported in [JUA13b]. Decomposition of four-level of EEG signals was done in order to obtain the frequency sub-bands of the EEG signals. Note that the number of scaling and wavelet coefficients is the same in all levels of MODWT. Figure 3.11 shows the delta (0-4 Hz) and alpha (8-12 Hz) sub-bands of an EEG signal of a Ictal subject, obtained using MODWT (Coif3). Plots on the left side of the figure correspond to sub-bands and plots on the right side correspond to their frequency components. Figure 3.12 and 3.13 show the same for an Interictal signal and for a Healthy signal, using MODWT(Coif5) and MODWT (Db6), respectively. Figures 3.11, 3.12 and 3.13 correspond to the decomposition of EEG signals using the criterion 1; a similar process was done for the criterion 2.

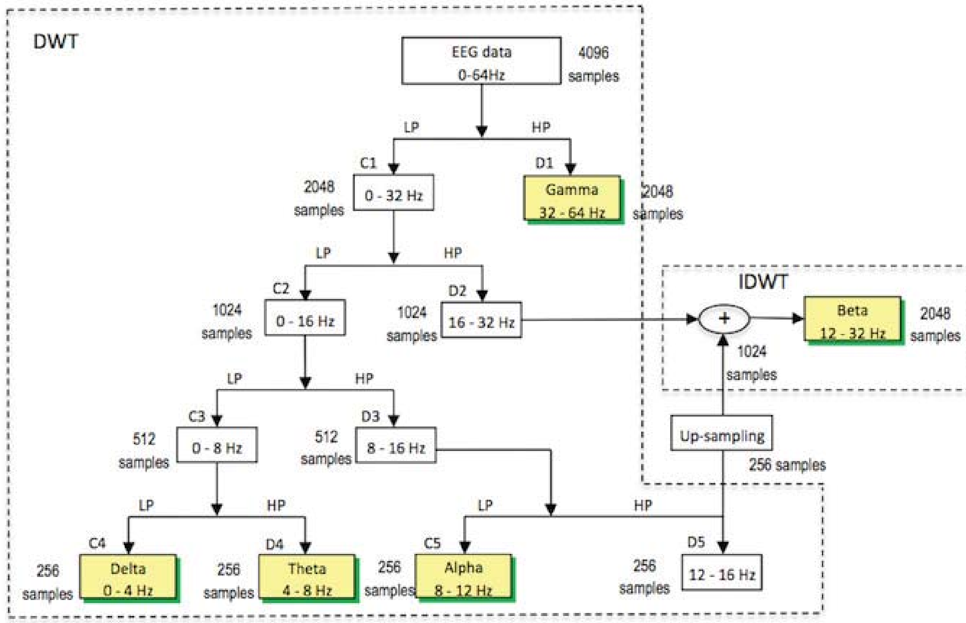

Figure 3.9: Decomposition of EEG in physiological sub-bands by DWT. This figure shows the name of sub-bands and its respective frequency ranges [JUA13b].

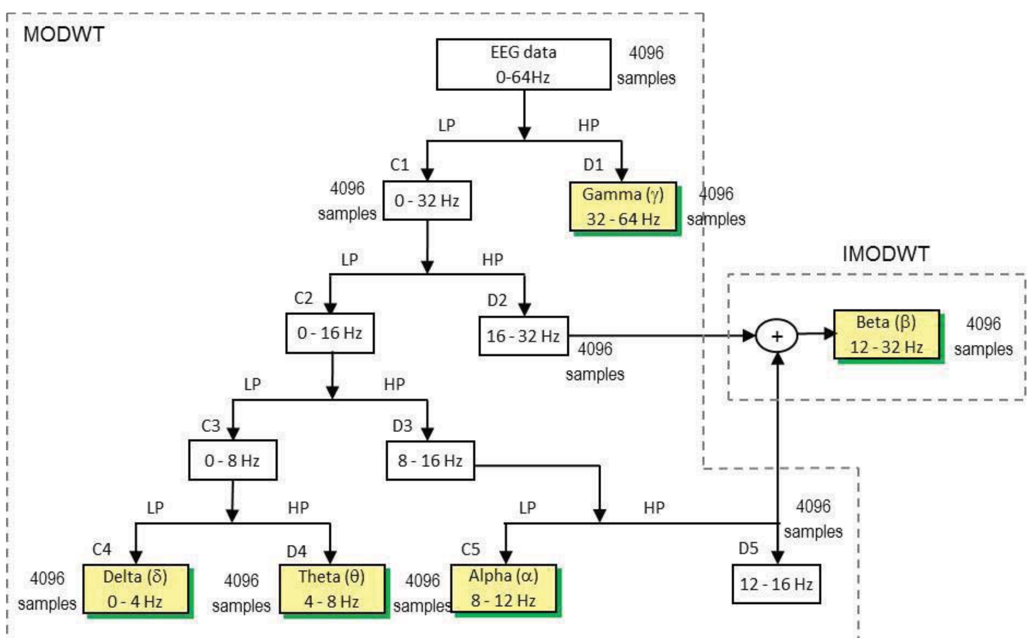

Figure 3.10: Decomposition of EEG in physiological sub-bands by MODWT. This figure shows the name of sub-bands and its respective frequency ranges [JUA13b].

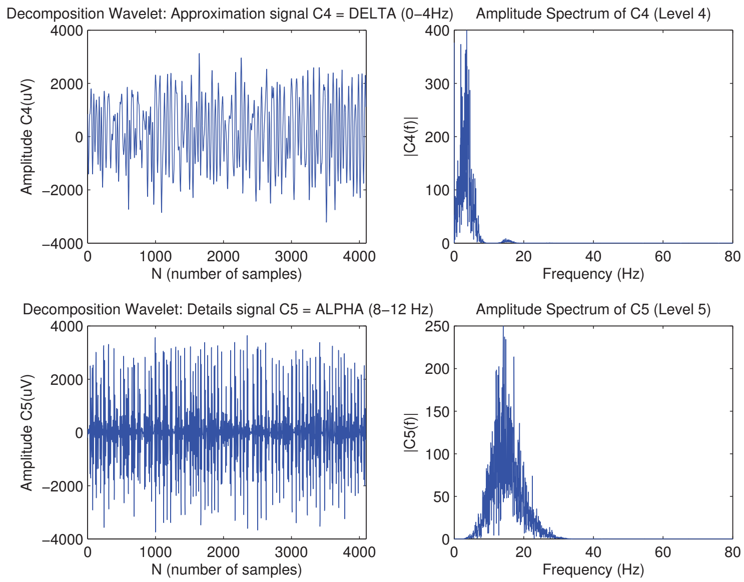

Figure 3.11: Delta and Alpha sub-bands of an EEG signal by MODWT (Coif3) of an Ictal subject. The graphs on the left side show the obtained sub-bands and the graphs on the right side show its corresponding frequency spectrum.

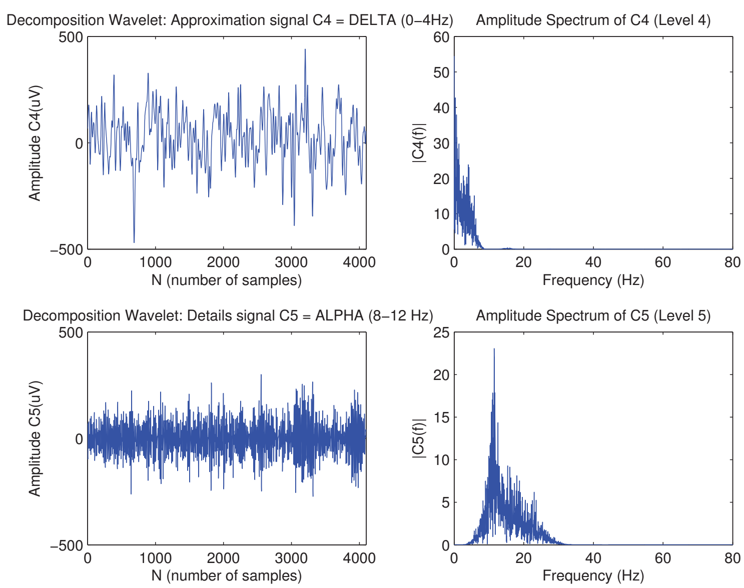

Figure 3.12: Delta and Alpha sub-bands of an EEG signal by MODWT (Coif5) of an Interictal subject. The graphs on the left side show the obtained sub-bands and the graphs on the right side show its corresponding frequency spectrum.

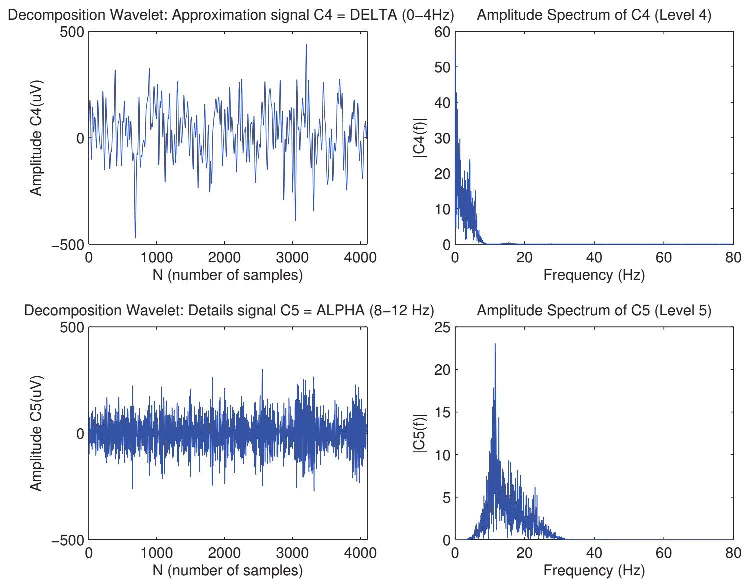

Figure 3.13: Delta and Alpha sub-bands of an EEG signal obtained by MODWT (Db6) of a Healthy subject. The graphs on the left side show the obtained sub-bands and the graphs on the right side show its corresponding frequency spectrum.

3.6 Multi-class Classification of EEG Signals.

Many classification tasks in real-world applications involve more than two classes, the so-called multi-class problems. Their application domain are diverse, for instance, in the field of bioinformatics, classification of microarrays and tissues, which operate with several class labels. Computer vision multi-classification techniques play a key role within objects, fingerprints and sign language recognition task, whereas in medicine, multiple categories are considered in problems such as cancer or electroencephalogram signals classification [GAL11].

Usually, it is easier to build a classifier to distinguish only between two classes than to build a classifier that distinguishes more than two classes in a problem. This is why binarization techniques have come up to deal with multi-class problems by dividing the original problem into easier to solve binary classification problems that are faced by binary classifiers. These classifiers are usually referred to as base learners or base classifiers of the system [GAL11].

Different methods to combine the outputs of the base classifiers have been developed; examples of these techniques include probability estimations, binary-tree based strategies, dynamic classification schemes and methods using preference relations. Other well-known combination methods are Pair-wise Coupling, Max-Wins rule or Weighted Voting [GAL11].

In this thesis, we proposed the use of binary-tree and one-vs-one (OVO) decomposition strategies with two aggregation methods: *Voting strategy* and *Weighted voting strategy*. These strategies are used in conjunction with our proposed method, called MRW-FFWNN (see Section 4.1), which distinguishes only between two classes. Next, we briefly describe these strategies.

Decision tree strategy. This learning strategy is one of the most widely used methods for inductive inference. It is a method for approximating discrete-valued target functions, in which the learned function is represented by a decision tree. Learned trees can also be represented as sets of if-then rules to improve human readability. These learning methods have been successfully applied to a broad range of tasks from learning to diagnose medical to learning to assess credit risk of loan applicants [MIT97].

Decision trees classify instances by sorting them down the tree from its root to some leaf node, which provides the classification of such instance. Each node in the tree specifies a test of some attribute of the instance, and each branch descending from that node corresponds to one of the possible values for this attribute. An instance is classified by starting at the root node of the tree, testing the attribute specified by this node, then moving down the tree branch corresponding to the value of the attribute in the given example. This process is then repeated for the subtree rooted at the new node [MIT97].

In this thesis, we have used a decision tree, but instead of using attributes in each node, we use a binary classifier, which decides the type of EEG signal. If an

EEG is not classified in any of the three classes is classified as indeterminate. We have tested several structures with binary-trees using different order in the questions related to each possible class. Each question is solved by a binary classifier, trained to distinguish between elements belonging to that class from the ones do not belonging to such class. The possible orders are:

- Ictal-Interictal-Healthy (Ic-In-H)
- Ictal-Healthy-Interictal (Ic-H-In)
- Interictal-Ictal-Healthy (In-Ic-H)
- Interictal-Healthy-Ictal (In-H-Ic)
- Healthy-Ictal-Interictal (H-Ic-In)
- Healthy-Interictal-Ictal (H-In-Ic)

Figure 3.14 shows a structure of the binary-tree, using the order Interictal-Healthy-Ictal (In-H-Ic).

One-vs-one decomposition scheme. OVO decomposition scheme divides a m class problem into $m(m-1)/2$ binary problems. Each problem is faced by a binary classifier, which is responsible for distinguishing between a different pair of classes. The learning phase of the classifiers is done using as training data a subset of instances from the original training data set. Such subset contains any of the two corresponding class labels, whereas the instances with different class labels are ignored [GAL11].

In recognition phase, a pattern is presented to each one of the binary classifiers. The output of a classifier given by $r_{ij} \in [0,1]$ is the confidence of the binary classifier discriminating classes $i = 1, ..., m$ and $j = 1, ..., m$ in favour of the former class, where m is the classes number. The confidence of the classifier for the latter is computed as $r_{ji} = 1 - r_{ij}$ in case that the classifier does not provide it (the class

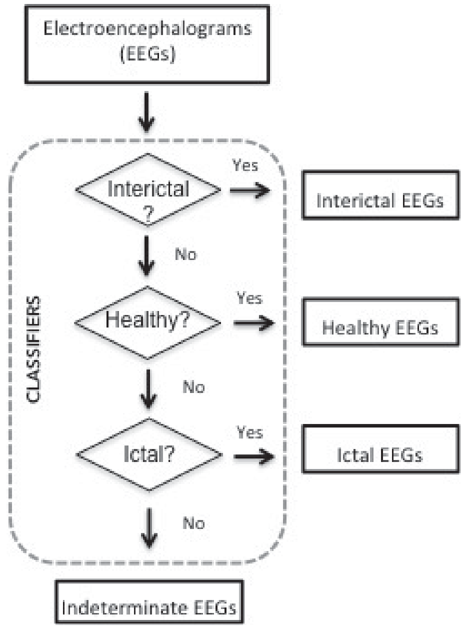

Figure 3.14: Binary-tree using the evaluation order: Interictal-Healthy-Ictal (In-H-Ic).

with the largest confidence is the output class of a classifier). These outputs are represented by a score matrix $R_{m\times m}$:

$$R_{m\times m} = \begin{bmatrix} - & r_{12} & \dots & r_{1m} \\ r_{21} & - & \dots & r_{2m} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \dots & - \end{bmatrix} \tag{3.1}$$

The final assigned class may be derived from the score matrix by different aggregation models. The aggregation methods that we used in this work are described below.

- *Voting strategy (VOTE):* Each binary classifier gives a vote for the predicted class. The votes received by each class are counted and the class with the largest number of votes is predicted:

$$Class = \arg\max_{i=1,\dots,m} \sum_{1<=j\neq i>=m} s_{ij} \tag{3.2}$$

where s_{ij} is the vote assigned by each binary classifier. Therefore, s_{ij} is 1 if $r_{ij} > r_{ji}$ and 0 otherwise. This method is also called binary voting or Max-Wins rule.

- *Weighted voting strategy (WV):* Each binary classifier votes for both classes. The weight for the vote is given by the confidence of the classifier predicting the class. The class with the largest sum value is the final output class:

$$Class = \arg \max_{i=1,\ldots,m} \sum_{1<=j\neq i>=m} r_{ij} \tag{3.3}$$

  where r_{ij} is the confidence of the binary classifier discriminating classes i and j in favour of the former class.

3.7 Discussion.

In this chapter, the description of proposed model for classification of three classes of EEG signals was presented. Here, we explained the three modules of the general block diagram of the proposed approach.

The EEG signals used in the experiments reported in this work come from a free-available EEG database, provided by the University of Bonn. Two approaches for filtering were tested, IIR and FIR filters. The filtering of the EEG signals was performed in order to remove noise added during recording in the module of preprocessing. The decomposition of EEG signals into delta (δ) and alpha (α) sub-bands and feature extraction were carried out using the DWT and MODWT by two criteria for the suitable wavelet choice. The first uses the wavelet chosen that gives the average highest of the correlations coefficients for each class of the EEG signals. The second criterion uses the wavelet that provides the highest number of times with the best correlation coefficient obtained for each class of the EEG signals. Each EEG signal was represented by a feature vector of six components, built using the mean, absolute median and variance of both delta (δ) and alpha (α) sub-bands. The feature vectors obtained are considered as inputs for the classifiers described in the next chapter, including our proposed classifier. Finally, in this chapter was presented, the description of the classification strategies used in this research: Binary-tree and OVO decomposition based on our proposed classifier MRW-FFWNN.