Automated Analysis of Epileptic EEG Using a Toolbox

María J. Angulo and Luis A. Rivera
Department of Electronics, Mechatronics and Biomedical Engineering
Universidad Del Valle de Guatemala
Guatemala, Guatemala
(ang16508, larivera) @uvg.edu.gt

***Abstract*—It is estimated that there are 50 million people affected by epilepsy worldwide. However, the study of this disease is still reserved for a reduced number of neurologists. Automatic detection of epileptic seizures can considerably improve the patients' quality of life. Current Electroencephalogram (EEG)-based seizure detection systems encounter defiances in real-life situations. Machine learning approach, support vector machines (SVM) and neural networks (ANN) may offered high performance in two-class EEG Classification. However, implementing this classifiers and improving their accuracy it's a real challenge. To address these challenge, we introduce the use of a toolbox that automatically extracts EEG features of epileptic and non epileptic seizures and also classifies the data using SVM and ANN algorithms. Specifically, to reveal ....**

***Index Terms*—Electroencephalogram (EEG), Epilepsy, Seizure detection, Toolbox, Machine Learning Classifiers.**

I. INTRODUCTION

Epilepsy is a chronic neurological disorder of the brain that affects people of all ages. It causes disorders of brain activity leading to clinical events called epileptic seizures. It is currently estimated that more than 50 million people worldwide suffer from this pathology [1], making it the second most common neurological diseases after migraine [2]. The diagnosis and treatment of epilepsy is an area of current research that has been expanded and improved by technological advances and the implementation of new algorithms for disease pattern recognition and seizure prediction.

A vast number of methods have been developed for automatic seizure detection using EEG signals. Extracting features that best describe the behaviour of EEGs is of great importance for automatic seizure detection systems' performance. Several feature extraction and selection techniques have been reported in the literature. Most of them use features in the time-domain, frequency-domain, time-frequency domain or sometimes in a combination of two domains.

This work introduces an automated toolbox to analyse epileptic EEG and to detect relevant patterns in the EEG signals of patients with epilepsy, using algorithms based on pattern recognition with machine learning approach, support vector machines (SVM) and artificial neural networks (ANN). The toolbox for MATLAB provides direct access to HUMANA database, which contains biomedical signals of epileptic patients from Epilepsy and Functional Neurosurgery Center (HUMANA) in Guatemala. It also allows users to load public domain EDF files with EEG data.

The Epileptic EEG Analysis Toolbox also provides functions useful for feature extraction and machine learning algorithms implementation. The Toolbox also contains a graphic interface that allows users to obtain classifications results quicker and easier. Thus this toolbox and also the HUMANA database provide a vital resource for MATLAB users who want to validate, and compare different automatic seizure detection systems. It is also intended for medics who wish to investigate and facilitate the diagnostic process of epilepsy.

II. RELATED WORK

The problem of EEG epileptic seizure detection as two-class classification is to differentiate between two distinct classes; Normal and Ictal EEG patterns [3].

A. Two-class EEG Classification

Most of the two-class seizure detection problems focus on the classification between normal EEG segments taken from healthy persons and seizure EEG patterns taken from epileptic patients while experiencing active seizures [3],[4] y [5]. Polat et al. achieved a higher classification accuracy of 98.68% using a decision tree (DT) classifier [4]. Wavelet transform was also used in [6] to analyze the EEG signals into five approximation and detail sub-bands. Then, the wavelet coefficients located in the low frequency range of $0-32$ Hz were used to compute the EEG features of energy and normalized coefficients. The linear discriminant analysis (LDA) classifier was used to prove the potential of the extracted features in detecting seizure onsets with a classification accuracy of 91.80%.

In this study SVM and MLP classifiers were used, extracting time-domain features from EEG signals. Wavelet transform was also used to analyze EEG signals into five sub-bands which cover the frequency bands of the five brain waves; beta, alpha, theta, delta and gamma. Then, all wavelet coefficients were used to compute the feature vector to implement the SVM and MLP classifiers.

B. EEG Feature Extraction

Choosing suitable features that can best represent the EEG signals is important to enhance EEG classification accuracy

[7]. Time domain features,like Mean Absolute Value (MAV), Zero Crossings (ZC), Slope Sign Changes (SSC) and standard deviation (std) and time-scale domain features using discrete wavelet transform (DWT) and sub bands decomposition are commonly used in several related works. Applying Discrete Wavelet Transform (DWT) on epilepsy-related EEG signal classification is gaining ground in recent years. The main advantage of DWT is that the resolution of time and frequency in DWT can be adapted to the frequency content of the examined patterns, thus leading to an optimal time-frequency resolution across all frequency ranges [8].

In this work, a combination of time domain and time-scale domain features are used. Implementing the same metodology as [9]. Angulo proposed two different feature extraction methods. The first one called "direct method", extracts time domain features: zero crossing, mean absolute value, standard deviation and kurtosis to pre-processed (filtered) EEG segments. The second method called "wavelet method" uses DWT coefficients with 7 level decomposition to cover the five brain waves in EEG signal. From each sub band features are extracted.

III. Methods

MATLAB Version R2019a was used for the development of the Toolbox. Epileptic EEG Analysis Toolbox contains graphical interfaces that give users unfamiliar with the classification algorithms the option to use them more easily.

Machine learning approach is used for feature extraction and epileptic seizure detection. The toolbox classification process starts with EEG data acquisition. Epileptic EEG Analysis Toolbox provides direct access to HUMANA database which contains EEG data from epileptic patients, it also allows user's local EEG data in EDF file format.

After EEG data is selected, feature extraction is required. The toolbox provides functions to extract time domain and time-scale domain features with parameters given by the user and also generates a feature vector and classes vector for classification algorithms implementation. Feature extraction generates a MAT file format (MATLAB workspace) with feature vector, labels vector and relevant information of EEG signal.

To generate the classes vector the toolbox assumes that EEG data is balanced (50% non seizure active and 50% seizure active samples) if this it's not the case, it also allows the user to load a correct classes vector before classification. Epileptic EEG Analysis Toolbox includes two classifications algorithms; support vector machine and artificial neural network. An MAT file is generated with classification results and accuracy. The process of classification is shown in figure 1.

A. Interface Structure

Epileptic EEG Analysis Toolbox interface has six main windows. The main menu in first windows includes "analyze data" option and HUMANA database connection. Second window includes "data acquisition",this interface allows user to load EDF files or selected existent files in HUMANA database and "consult patients" option which enables third window, "Database Access". This interface provides functions to interact with patient records and database options. The fourth window interface "Feature Extraction" provides functions to generate features and labels vector (which includes binary classification classes). The fifth and sixth windows include "SVM" and "ANN" algorithms implementation and generate classification results. Interface structure is shown in figure 2.

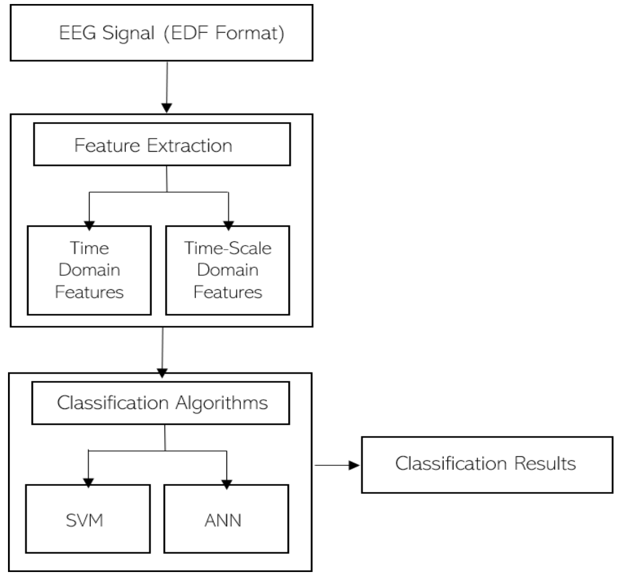

Fig. 1. Epileptic EEG Analysis Toolbox Framework.

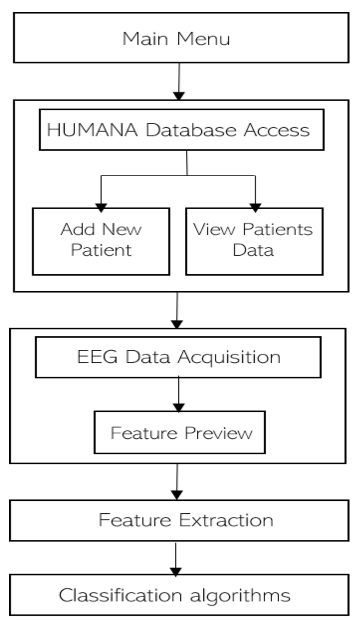

Fig. 2. Epileptic EEG Analysis Toolbox Interface Structure.

IV. Results

A. Data Selection

In this work, two different datasets were used to test the performance of the toolbox. The first dataset was from University of Bonn (UBonn) [10]. It had five sets denoted AE, each containing 100 single channel EEG segments of 23.6-sec duration with a sampling rate of 173.61 Hz. These segments were selected and cut out from continuous multichannel EEG recordings after visual inspection for artifacts. The scalp EEG signals were digitally filtered using a 48th-order FIR high-pass filter (hamming window) with the cutoff frequency at 0.5Hz. This database does not include EDF files. All EEG signals from the five sets are defined as TXT files.

The second dataset was collected at the Children's Hospital Boston, Massachusetts (MIT), consists of EEG recordings from pediatric subjects with intractable seizures in EDF file format. Subjects were monitored for up to several days following withdrawal of anti-seizure medication in order to characterize their seizures and assess their candidacy for surgical intervention [11]. The EEG signals were sampled at 256Hz and digitally filtered by a 48th-order FIR high-pass filter (hamming window) with the cutoff frequency at 0.5Hz to remove low-frequency artifacts. The 256Hz sampling rate is large enough to cover general human EEG rhythms (bandwidths), including, (¡ 4Hz), (4 7Hz), (8 15Hz), (16 31Hz) and (¿ 31Hz).

From UBonn dataset, set A was used as normal patterns (non-seizure) and set E was treated as ictal (active seizure). In order to keep the data balance, in an MAT file set A with set E were concatenated into a data vector of which 50% of the elements were zeros representing the normal class (non-seizure) and 50% were ones representing the ictal class (active seizure).

The second dataset used, includes recordings grouped into 23 cases. Each case contains between 9 and 42 continuous EDF files from a single subject. In this work, 5 different cases were selected to test classification algorithms from the toolbox. From each case, recording segments were chosen to keep data balanced. An MAT file with recording segments of the same 5 cases was created (50% data have non seizure samples and 50% with seizures). Finally another case of Physionet data base was selected to test feature extraction methods from Epileptic EEG Anaylisis Toolbox.

B. Feature Results

In Figure a

C. Classification Results

Physionet MAT file SVM and ANN classification results are shown in figure 3. The confusion matrices as example shown in Figure 4 evaluates the classification algorithms performance using direct method for feature extraction (time domain features: mean absolute value, kurtosis, zero crossing and standard deviation) and wavelet method (time-scale domain features: power, skewness, standard deviation, kurtosis, mean and zero crossings). Results in this part showed a consistent pattern across high accuracy in direct method for feature extraction. Nevertheless, using suitable wavelet and decomposition level, provide promising seizure detection accuracy with wavelet method for feature extraction. The versatility of the toolbox allows the user to interact with this parameters and settings which improve the accuracy of the classification algorithms easily.

Dataset	Feature Extraction Method	Classification Algorithm Accuracy
Physionet	Direct ZC, mean, kurtosis and standard deviation	SVM : 99.5% ANN: 100%
UBonn	Direct ZC, mean, kurtosis and standard deviation	SVM: 100% ANN: 100%
Physionet	Wavelet power, mean, kurtosis, standard deviation, skewness and zc	ANN : 75.1%
UBonn	Wavelet	SVM: 97.5% ANN: 99.1%

Fig. 3. Accuracy Results: SVM and ANN for Physionet/UBonn MAT Files

Fig. 4. Confusion Matrices Example: ANN Results for Physionet MAT File

A different physionet case is shown in Figure . This example shows the number of zero crossings of samples with seizure in 1 s windows (256 samples).

Epileptic EEG Analysis Toolbox, provides a vital resource for MATLAB and users who want to learn, validate, and compare different feature extraction methods and machine learning classification algorithms to analyze epileptic EEG data with a graphic interface which simplify the data analysis.

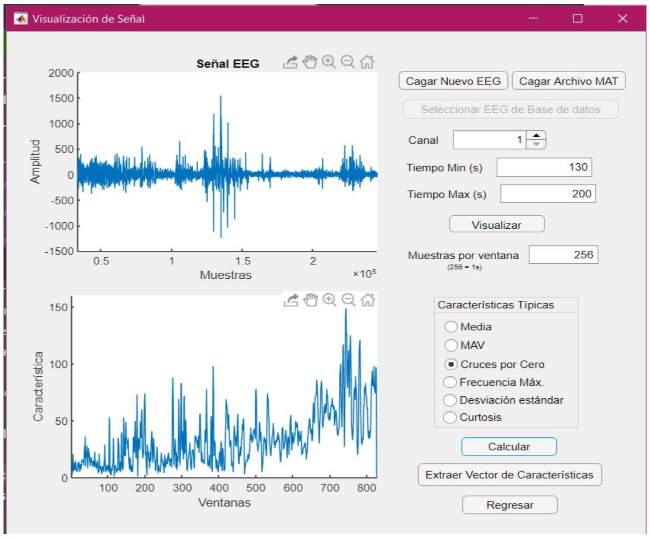

Fig. 5. Feature Extraction Example: ZC for Physionet Case Chb02-16

It also allows the users to explore the optimal combination of factors such as window size, time domain feature combination, number of channels, time-scale domain feature combination, mother wavelet, decomposition level, frequency band, etc. Efficiently setting some of these factors will lead to high seizure detection accuracy.

References

[1] (Jun. 2019). "Notas descriptivas: Epilepsia," [Online]. Available: https://www.who.int/es/news-room/fact-sheets/detail/epilepsy.

[2] G. Rogers, "R. appleton and a. marson 2009: Epilepsy: The facts, oxford, uk: Oxford university press. 186 pp,£ 9.99 paperback. isbn: 9780199233687," *Primary Health Care Research & Development*, vol. 11, no. 4, pp. 413–413, 2010.

[3] A. Aarabi, F. Wallois, and R. Grebe, "Automated neonatal seizure detection: A multistage classification system through feature selection based on relevance and redundancy analysis," *Clinical Neurophysiology*, vol. 117, no. 2, pp. 328–340, 2006.

[4] K. Polat and S. Güneş, "Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 1017–1026, 2007.

[5] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines," *Expert Systems with Applications*, vol. 39, no. 1, pp. 202–209, 2012.

[6] Y. U. Khan, N. Rafiuddin, and O. Farooq, "Automated seizure detection in scalp eeg using multiple wavelet scales," in *2012 IEEE international conference on signal processing, computing and control*, IEEE, 2012, pp. 1–5.

[7] L. Guo, D. Rivero, and A. Pazos, "Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks," *Journal of neuroscience methods*, vol. 193, no. 1, pp. 156–163, 2010.

[8] E. D. Übeyli, "Analysis of eeg signals by combining eigenvector methods and multiclass support vector machines," *Computers in Biology and Medicine*, vol. 38, no. 1, pp. 14–22, 2008.

[9] M. J. Angulo, "Análisis y reconocimiento de patrones de señales biomédicas de pacientes con epilepsia.," Universidad del Valle de Guatemala, Tech. Rep., 2020.

[10] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, p. 061 907, 2001.

[11] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Institute of Technology, 2009.