Multi-class EEG Signal Classification for Epileptic Seizure Diagnosis

Dalila Cherifi[1(✉)], Laid Afoun[1], Zakaria Iloul[1], Billal Boukerma[1], Chaouki Adjerid[1], Larbi Boubchir[2], and Amine Nait-Ali[3]

[1] Institute of Electrical and Electronic Engineering, University of Boumerdes, Boumerdes, Algeria
dalilache2013@gmail.com
[2] LIASD-University of Paris 8, Paris, France
[3] LiSSi, EA 3956, Paris-Est Creteil (UPEC), Paris, France

Abstract. EEG signal recordings are increasingly replacing the old methods of diagnosis in medical field of many neurological disorders. Our contribution in this article is the study and development of EEG signal classification algorithms for epilepsy diagnosis using one rhythm; for classification, an optimum classifier is proposed with only when used one rhythm so that both execution time and number of features are reduced. We used wavelet packet decomposition (WPD) to extract the five rhythms of brain activity from the public Epilepsy-EEG recordings in order to represent each signal with features vector; then we applied on it the well-known classification methods. A statistical study is done to validate the different algorithms.

Keywords: EEG · Wavelet packet decomposition · Features extraction · Epilepsy diagnosis

1 Introduction

The internet and digitization revolution had led to the presence of huge amount of data in hand, whether the data is previous recordings of forecast in a certain area or stock market prices over few previous hours. The medical field had his own benefits from this revolution since the previous situation of the patient (age, weight, height and blood pressure.) all is recorded in computers data base which doctors can access and make much more accurate diagnosis and more appropriate treatment, but still some situations where diagnosis is difficult or make predictions about disease evolution is not really accurate because even the existence of huge data is hard to analyze by human being and extract some useful information. This was the situation until artificial intelligence and machine learning algorithms and recently deep learning took the role of data analyst away since it is more accurate and less time consuming. Electroencephalography (EEG) records the brain's activities by measuring the voltage fluctuation on the scalp; although, the signals recorded can be influenced by mood, stress, as well as the mental state of the individual being tested. Scalp EEG activity shows oscillations at a variety of frequencies, mainly in the [1, 40] Hz range and can be classified using a set of features like Autoregression [1, 2], wavelet transform [3, 4], Entropy [5], Energy

M. Hatti (Ed.): ICAIRES 2020, LNNS 174, pp. 635–645, 2021.
https://doi.org/10.1007/978-3-030-63846-7_60

Spectrum Density, spatial filter [6]. There are many researches about neurological diseases in general and epilepsy in particular and related machine learning and EEG applications, but they take all the signal, this needs a long time execution and large features vector size.

The aim of this work is to study and develop EEG signal classification algorithms for epilepsy diagnosis by using machine learning and examining only one rhythm. To carry out this work, the article has been divided into four parts, briefly described below: The Sect. 1 introduces a general concepts about brain activity, acquisition systems to detect brain disorders and some familiar neurological diseases such as epilepsy. Section 2 lists some signal features that will be used in this work, and then describes some well-known classification methods. The Sect. 4 explores the experimental investigations of this study, statistical analysis for various methods proposed as well as their performances.

2 EEG Overview

The brain is the most complex part of the body, it is the part where all neurological actions take place; brain activity is the firing that occurs between neurons in response to certain situation, the brain receives messages from sensory systems as electrochemical signals that can treat and output a decision about what should be done. EEG is the abbreviation of “Electro-Encephalo-Gram”; it is a recording of the electrical changes occurring in the brain produced by electrodes placed on the scalp and amplifying the electrical potential developed (see Fig. 1) [7]. Typical EEG systems can have from as few as a single channel to as many as 256 channels. But using larger electrode arrays (dense array EEG) and expanding to more channel EEG systems provide several advantages [8, 9]. There are 5 rhythms (alpha, beta, gamma, delta and theta) that can be extracted from an EEG signal; they exist in the brain depending on the situation of the person.

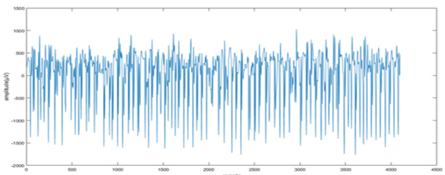

Fig. 1. One channel EEG signal.

3 EEG Based Methodology for Epilepsy Diagnosis

A. Epilepsy Disease

Epilepsy is a brain disorder that takes the form of seizures caused by sudden change in the brain electrical activity that may develop from imbalanced brain wiring, an imbalance in neurotransmitters production in cells or a combination of both. It is one of the most frequent neurological diseases experienced worldwide [10]. As result, behaviors the patient may have are: temporary loss of breath and memory loss, jerky movements and awareness. The number of seizures and the duration vary from one patient to another (Fig. 2).

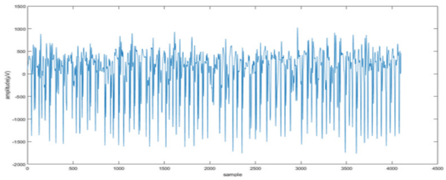

Fig. 2. EEG signal having epilepsy [10].

In this section, a biomedical system based on EEG signal for epilepsy diagnosis is developed (see Fig. 3). A pre-processing stage which reduces the signal noise by a low pass filter is first presented, then the needed features are extracted using wavelet packet decomposition as well as some statistical parameters (The mean, variance, standard deviation, first quartile, third quartile and inter quartile (IQR), median, mode). Finally a Classification based methods are tested followed by an experimental part.

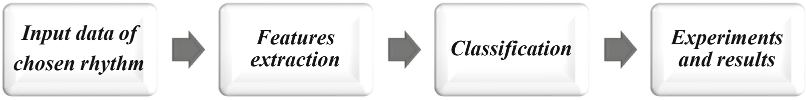

Fig. 3. Global diagram of the whole procedure.

B. Pre-Processing

The EEG is usually analyzed by physicians and medical specialists particularly for the detection of neural rhythms. Although, the signal is mainly affected by various sources of noise and mixed with other biological signals whose common artifact sources are the power line interference (50 or 60 Hz). To remove or correct the artifacts from the EEG signal without losing necessary data, a Butterworth filter has been adopted. Depending on the types of EEG noises, and taking into consideration the frequency bands of the five rhythms that contain the necessary features, this yields the use of a low-pass Butterworth filter with a frequency band of 60 Hz. The achievement of this filter is significant when it is applied to the EEG recorded signal, as shown in Fig. 4.

C. Features Extraction

Each signal can be represented by features that define its properties whether its time properties or frequency properties or energy properties [10]. After the pre-processing stage, a filtered EEG signal suitable for extracting the needed features was obtained.

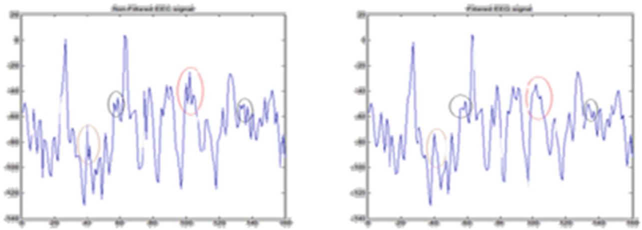

Fig. 4. Portion of an EEG signal before and after the filtering process.

In our work a method was used merging two kinds of feature extraction techniques, dealing with both the For different individuals, the energy distribution of the frequency components between individuals is quite different and this option makes it possible to adapt those frequency components as features to represent the EEG signals. Improvising from this fact, a 4-Level WPD is applied to each channel of the pre-processed EEG signal for each subject and only coefficients from nodes {(1, 1), (2, 1), (3, 1), (4, 0) and (4, 1)} representing the frequency bands of the main 5 EEG rhythms (Gamma, Beta, Alpha, Delta, and Theta) are extracted.

To form the feature vectors: the median, mode, mean, standard deviation, maximum, minimum, first quartile, third quartile and inter quartile were evaluated for each node of every channel (see Fig. 5). Since there are three statistical parameters and five nodes, there are 9 * 5 = 45 features obtained for each channel associated to every subject. Finally, the 64 channels data are stored, knowing that each data holds the median, mode, mean, standard deviation, maximum, minimum, first quartile, third quartile and inter quartile data of the five rhythms [11]. These features were stored for future use as data for the training part while testing.

- **The mean**
  It defines the average value of the data in hand, the value of the mean helps to identify where the data points are concentrated in classification, it is the sum of the numbers divided by how many numbers there are.

$$\mu = \sum_{k=1}^{N} \frac{x_k}{N} \tag{1}$$

  Where μ is the mean; x_k are the values in the data set; N is the number of elements.
- **The variance**
  Variance is represented by σ^2 and its mathematical expression is:

$$\sigma^2 = \frac{1}{N} \sum_{k=1}^{N} (x_k - \mu)^2 \tag{2}$$

- **The standard deviation**
  The standard deviation quantify the amount of dispersion in a data set or by how much the data set deviate from the mean and it is the square root of the variance

$$\boldsymbol{\sigma} = \sqrt{\frac{1}{\boldsymbol{N}}\sum\nolimits_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{N}} (\boldsymbol{x}_{\boldsymbol{k}} - \boldsymbol{\mu})^2} \tag{3}$$

  Where σ^2 is the variance, x_k are the values in the data set and μ is the mean.
- **First quartile, third quartile and inter quartile (IQR)**
  The first quartile is the median of the first half of the data set while third quartile is the median of the second half. The inter quartile is a measure of the dispersion of the data; it is the difference between first and third quartiles.
- **The median (second quartile)**
  The Median value is the middle number in a sorted list of data.
- **The mode**
  The mode is a statistical term that represents the most frequently occurring number found in a sorted list of data.

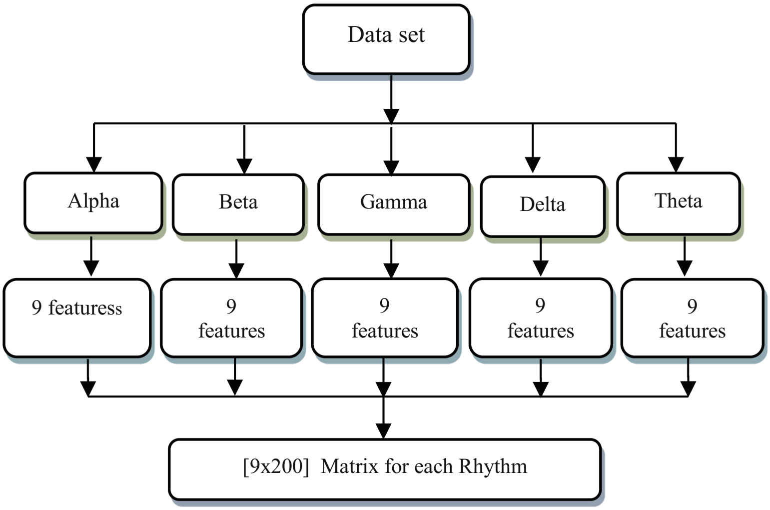

Fig. 5. Rhythms and features extraction from the data set.

D. Classification Algorithms

- **Decision Tree**
  Decision Tree Builds classification models in the form of tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is developed [12].

- **Quadratic Discriminant**
  Quadratic Discriminant analysis is a dimensionality reduction technique which aims to decrease the dimensions by eliminating the dependent and redundant features [13].
- **Logistic regression**
  Logistic regression estimates probabilities using a logistic/sigmoid function by evaluating the relationship between the categorical dependent variable and one or more independent variables by [14].
- **Support Vector Machine (SVM)**
  Support Vector Machine (SVM) can solve linear and non-linear problems and work well for many practical problems related to classification and regression [15].
- **K-Nearest Neighbor (K-NN)**
  K-Nearest Neighbor is a classification method based on finding the clusters; it computes the euclidian distance between the point and the mean of each cluster in order to form new means until no changes occur [16].
- **Ensemble**
  Ensemble methods are a combination of several classification methods that already exist in one predictive model. The purpose of that is to decrease the variance (bagging), decrease the bias (boosting) or improve prediction (stacking) [17].

4 Results and Discussion

In this part, we present the results and discussions of the classification methods on which we use features vector selected for the classification consists from: the median, mode, mean, standard deviation, maximum, minimum, first quartile, third quartile and inter quartile are used in this implementation. The purpose of the experiments is to identify which rhythm is mostly affected epilepsy and which classifier should be used.

A. The Data Set Description

The EEG dataset includes recordings for both healthy and epileptic subjects (see Table 1, 2). The dataset includes five subsets (denoted as A, B, C, D, and E) each containing 100 single channel EEG segments, each one having 23.6 s duration. The subsets A and B have been acquired using surface EEG recordings of five healthy volunteers with eyes closed and open, respectively. Signals in two sets have been measured in seizure-free intervals from five patients in the epileptogenic zone (set D) and from the hippocampus formation of the opposite hemisphere of the brain (set C). Finally, subset E contains seizure activity, selected from all recording sites exhibiting actual activity [10]. We have used the above-described dataset to create four different classification problems.

B. Experiments and Results

The purpose of the first part is to identify which rhythm is mostly affected by epilepsy and which classifier should be used. After answering those two questions, we use a band pass filter to extract only the rhythm in question in order to reduce time execution of WPD. This gives the optimum classifier: First, the EEG signals are decomposed to

Table 1. Data separation according to classification problem.

Classification problem	Classes	Number of segment
1	Normal (Set A- Set B) Seizure free (Set C- Set D) Seizure (Set E)	200 200 100
Total		**500**
2	Non Seizure (Set A- Set B- Set C- Set D) Seizure (Set E)	400 100
Total		**500**
3	Normal (Set A) Seizure free (Set C) Seizure (Set E)	100 100 100
Total		**300**
4	Normal (Set A) Seizure (Set E)	100 100
Total		**300**

Table 2. Detailed description of the data sets.

	Set A	Set B	Set C	Set D	Set E
Subjects	Five healthy subjects	Five healthy subjects	Five epileptic subjects	Five epileptic subjects	Five epileptic subjects
Patient's state	Eyes open	Closed eyes	Seizure free	Seizure free	Seizure activity
Type of the electrode	Surface	Surface	Intracranial	Intracranial	Intracranial
Placement of the electrode	International 10–20 system	International 10–20 system	Opposite to: Epileptogenic	Within Epileptogenic	Within Epileptogenic
Number of channel	100	100	100	100	100
Time duration (s)	23.6	23.6	23.6	23.6	23.6

extract five rhythms; for each rhythm, the features vector selected for the classification is of dimension [1 $\times$ 9]; this gives us a matrix size [9 $\times$ 200] for each rhythm.

The program is decomposed into four steps:

- Filtering the data: chosen belongs to the fourth class is filtered using Butterworth filter after sampling.
- The wavelet decomposition algorithm is used to extract the five rhythms that formulate the EEG signals. For each rhythm, nine features are extracted; the mean, the standard deviation, the minimum, the maximum, the median and the mode, first quartile, third quartile and inter-quartile.

- In order to use classification learner app in MATLAB, a table should be constructed; this table contains the values of the features for each rhythm alone added to that a description of whether the signal is for a healthy or epileptic person for all data set.
- A cross validation procedure is used which is a modal validation technique that partition the data into training subsets and validation (testing) subsets to reduce variability; each time, one subsets used for testing nine subsets is used for training. The performance of the system is mesured by averaging the testing results. The classification app gives the accuracy of many classification algorithms. The accuracy is the number of correct decisions divided by the total number of cases; the accuracy is within ±0.5% error, the results are shown in tables below (Tables 3, 4, 5, 6, 7 and 8):

Table 3. Alpha rhythm

	E-A	E-B	E-D	E-C	Mean
Simple tree	99.5%	94.5%	97.5%	96.5%	97%
Quadratic discriminant	99.5%	91.5%	95.5%	98.0%	96.13%
Logistic regression	99.5%	95.5%	96.0%	97.5%	97.13%
Fine Gaussian SVM	99.5%	97.0%	96.0%	98.0%	97.63%
Fine KNN	99.5%	96.5%	94.5%	99.0%	97.38%
Ensemble bagged trees	99.5%	96.0%	96.5%	98.0%	97.5%

C. Discussion

We see from the above table that the classification using alpha rhythm gives the best results; we see also that the algorithm with the best accuracy is fine Gaussian SVM for the same alpha rhythm. Because WPD takes a long time we don't need to calculate all rhythms, we can take just one rhythm using band-pass filter to reduce the Running time. As in the previous part, the data set is filtered but instead of using WPD we used a band pass filter to extract just the range of alpha rhythm, this will decrease the time execution. An SVM model is trained by the data proposed, the training data includes sets A and E from 1 to 90 so that the last 10 channels of each set can be used for testing. We noticed that using the filter reduces time execution compared to WPD. The last point in classification is that we trained the SVM model and tested it to see the accuracy after using a single rhythm with only three features which are the mean, the minimum and maximum. We found that these features are the best to describe the peaks of the epileptic EEG.

Table 4. Beta rhythm

	E-A	E-B	E-D	E-C	Mean
Simple tree	98.5%	95.5%	97.5%	98.0%	97.38%
Quadratic discriminant	99.0%	95.5%	94.5%	97.5%	96.63%
Logistic regression	98.5%	95.5%	98.0%	98.0%	97.5%
Fine Gaussian SVM	99.5%	96.5%	94.5%	97.5%	97.0%
Fine KNN	99.5%	95.0%	92.5%	96.5%	95.88%
Ensemble bagged trees	98.5%	96.5%	95.5%	98.0%	97.13%

Table 5. Gamma rhythm

	E-A	E-B	E-D	E-C	Mean
Simple tree	97.0%	96.0%	97.0%	98.0%	97.0%
Quadratic discriminant	99.0%	95.5%	92.5%	98.0%	96.25%
Logistic regression	99.5%	96.0%	98.5%	95.0%	97.25%
Fine Gaussian SVM	99.5%	96.0%	95.0%	98.0%	97.13%
Fine KNN	99.5%	95.5%	91.5%	96.5%	95.75%
Ensemble bagged trees	98.5%	96.5%	96.0%	97.5%	97.13%

Table 6. Delta rhythm

	E-A	E-B	E-D	E-C	Mean
Simple tree	99.0%	98.0%	89.5%	95.0%	95.38%
Quadratic discriminant	100.0%	99.0%	84.5%	93.5%	94.25%
Logistic regression	99.0%	99.0%	92.5%	96.5%	96.75%
Fine Gaussian SVM	99.5%	99.0%	89.5%	94.5%	95.63%
Fine KNN	99.5%	99.0%	90.0%	93.0%	95.38%
Ensemble bagged trees	99.5%	99.0%	92.0%	95.0%	96.38%

Table 7. Theta rhythm

	E-A	E-B	E-D	E-C	Mean
Simple tree	99.5%	93.0%	96.5%	98.0%	96.75%
Quadratic discriminant	99.5%	91.0%	93.0%	98.5%	95.5%
Logistic regression	99.0%	95.0%	97.0%	96.5%	96.88%
Fine Gaussian SVM	99.5%	93.5%	95.5%	98.5%	96.75%
Fine KNN	99.5%	92.0%	92.5%	98.5%	95.63%
Ensemble bagged trees	99.5%	94.0%	97.0%	98.0%	97.13%

Table 8. Rythms affected by epilepsy

Alpha rhythm	Beta rhythm	Delta rhythm	Gama rhythm	Theta rhythm
97.13%	96.92%	95.63%	96.75%	96.44%

5 Conclusion

Brain activity is the response of the brain and neuronal system to exterior excitation, Electroencephalography (EEG) signal is a recording of this activity in certain duration of time, the EEG or brain activity signal is useful tool for neurologists to see the effect of brain disorders on its five main rhythms (alpha, beta, gamma, delta and theta). From the statistical study that we conducted, we deduced that alpha is mostly affected by epilepsy disorder and that the best classification algorithm that can help neurologists to diagnose this disorder is Support Vector machine (SVM) among many others since it gives more accuracy. After finding the rhythm, a band pass filter is used to extract only its band to use it in training the model instead of using the Wavelet Packet Decomposition (WPD), this reduced the time execution since it extracts only the needed rhythm. The number of features also reduced to three (mean, minimum, maximum) since they describe well the differences between healthy and seizure EEG signals.

References

1. Anderson, W., Stolz, E.A., Shams, S.: Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. IEEE Trans. Biomed. Eng. **45**(3), 277–286 (1998)
2. McFarland, J., Wolpaw, J.R.: Sensor motor rhythm-based brain–computer interface (BCI): model order selection for autoregressive spectral analysis. J. Neural Eng. **5**(2), 155–162 (2008)
3. Amin, H.U., Malik, A.S., Ahmad, R.F., Badruddin, N., Kamel, N., Hussain, M., Chooi, W.T.: Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australas. Phys. Eng. Sci. Med. **38**(1), 139–149 (2015)
4. Jahankhani, P., Kodogiannis, V., Revett, K.: EEG signal classification using wavelet feature extraction and neural networks. In: IEEE John Vincent Atanasoff, International Symposium on Modern Computing, Sofia, pp. 120–124 (2006)
5. Sabeti, M., Katebi, S., Boostani, R.: Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. Artif. Intell. Med. **47**(3), 263–274 (2009)
6. Nandish, M., Stafford, M., Kumar, P.H., Ahmed, F.: Feature extraction and classification of EEG signal using neural network based techniques. Int. J. Eng. Innov. Technol. (IJEIT) **2**(4), 1–5 (2012)
7. Tatum IV, W.O.: Handbook of EEG Interpretation. Demos Medical Publishing (2014)
8. NeuroSky: Multi-channel EEG (BCI) devices. http://neurosky.com/2015/07/multi-channel-eeg-bci-devices/

9. Ting, W., Guo-Zheng, Y., Bang-Hua, Y., Hong, S.: EEG feature extraction based on wavelet packet decomposition for brain computer interface. Measur. J. Int. Measur. Confederation **41** (6), 618–625 (2008)
10. https://www.physionet.org/physiobank/database/eegmmidb/
11. Siuly, S., Lin, Y., Zhang, Y.: EEG Signal Analysis and Classification Techniques and Applications. Springer (2016)
12. Kamiński, B., Jakubczyk, M., Szufel, P.: A framework for sensitivity analysis of decision trees. Cent. Eur. J. Oper. Res. **26**(1), 135–159 (2017)
13. Tharwat, A., Gaber, T., Ibrahim, A., et al.: Linear discriminant analysis: a detailed tutorial. AI Commun. **30**(2), 169–190 (2017)
14. Hosmer, D.W., Lemeshow, S.: Applied Logistic Regression. Wiley Series in Probability and Mathematical Statistics, New York (2000)
15. Schölkopf, B., Smola, A.J., Bach, F.: Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge (2002)
16. Coomans, D., Massart, D.L.: Alternative k-nearest neighbour rules in supervised pattern recognition: part 1. k-nearest neighbour classification by using alternative voting rules. Anal. Chim. Acta **136**, 15–27 (1982)
17. Rokach, L.: Ensemble-based classifiers. Artif. Intell. Rev. **33**(1–2), 1–39 (2010)