Epileptik ve Normal EEG Sinyallerinin Alt Bant Güç Spektrumu Kullanılarak Sınıflandırılması Classification of Epileptic and Normal EEG Signals Using Power Spectrum of Sub-bands

Sude Pehlivan[1], Savas Sahin[2]
[1]Department of Biomedical Technologies, Izmir Katip Celebi University, Izmir, Turkey
sudepehlivan35@gmail.com
[2]Department of Electrical and Electronics Engineering, Izmir Katip Celebi University, Izmir, Turkey
phd.savas.sahin@gmail.com

***Özetçe*—Dünya genelinde birçok insanın hayatını etkileyen epilepsinin erken teşhisi, hastaların hayatlarına verimli devam edebilmesi için uygulanacak tedavinin ilk adımıdır. Uzmanlar, bu teşhisin en kısa sürede ve en doğru şekilde yapılması için çok fazla zaman ve enerji harcamak zorunda kalmaktadır. Bu çalışmanın amacı, nöbetleri otomatik olarak teşhis edebilen bir sistem geliştirmek için makine öğrenmesi algoritmalarının epileptik ve normal sinyalleri ayırt etme kapasitesini araştırmaktır. LabVIEW, hem epileptik hem normal kayıtlar için bir öznitelik olarak kullanılan EEG alt bant güçlerinin toplamını bulmak için kullanılmıştır. Bu öznitelikler Matlab kullanılarak farklı sınıflandırıcılar ile sınıflandırılmış ve sınıflandırma sonucunda alt bant güç toplamının epileptik ve normal EEG sinyallerinin sınıflandırılmasında anlamlı bir öznitelik olarak kullanılabileceği sonucuna varılmıştır.**

Anahtar Kelimeler*—LabVIEW; Epilepsi; Makine Öğrenmesi.

***Abstract*—The early diagnosis of epilepsy, which affects the lives of many people worldwide, is the first step of treatment to help patients to continue their lives efficiently. Experts have to spend a lot of time and energy to make this diagnosis as quickly and accurately as possible. The aim of this study was to investigate the capacity of machine learning algorithms to distinguish epileptic and normal signals to develop a system that can automatically diagnose seizures. LabVIEW was used to obtain the sum of EEG sub-band powers which were used as an attribute for both epileptic and normal records. These attributes were classified with different classifiers using Matlab and as a result of the classification, it was concluded that the sub-band power sum can be used as a meaningful attribute in the classification of epileptic and normal EEG signals.**

Keywords*—LabVIEW; Epilepsy; Machine Learning.

I. INTRODUCTION

Brain, the control center of the body, is formed by cells called neurons that are receiving and transmitting information in the form of electricity. Electroencephalogram (EEG) is a device that can measure the electrical activity of the brain. Due to its easy usage, costless, and reliable procedure, it is widely used for the diagnosis of several neurological and physiological diseases [1]. Epilepsy is an abnormality of the brain which is caused by undesired discharges of the neurons which can be detected via EEG using electrodes [2]. There exist several signal processing techniques that extract meaningful features from different signal types. They consist of frequency domain analysis such as Fourier Transform (FT), time-domain analysis such as statistical analysis, time-frequency analysis such as Wavelet Transform (WT) and non-linear measures [3].

Power Spectrum analysis is based on FT which is one of the most used feature extraction techniques and widely used. Chua et al. conducted a comparative study for the performance of Higher-Order Statistics (HOS) based features and power spectrum based features for the classification of normal, perictal, and epileptic EEG signals. Mean of spectral magnitude for PSD, mean of spectral magnitude for HOS, entropy for Power Spectral Density (PSD), and entropy for HOS were calculated as features. The Gaussian Mixture Model (GMM) was selected to determine the performance of two different feature extraction methods. Accuracy of HOS based features was 93.11% while the accuracy of PSD based features was 88.78%. It was concluded that during the seizure, entropy values were increased because of spontaneous discharge of neurons, and HOS based features outperformed PSD features [4].

II. MATERIALS AND METHOD

A. Dataset

In this study, an open-source EEG dataset described in the study of Andrzejak et al. was used [5]. It consists of 5 sets as A, B, C, D, and E. Dataset A includes recordings from healthy volunteers as eyes open, B includes healthy recordings as eyes closed, sets C and D are seizure-free, and set E includes seizure activity. A and B sets were recorded with extracranial electrodes and set E was recorded with intracranial electrodes. 128 channel amplifier set was used for recording and the sampling frequency was 173.61 per second. In this study, sets A and E were used to differentiate healthy volunteers and epileptic patients.

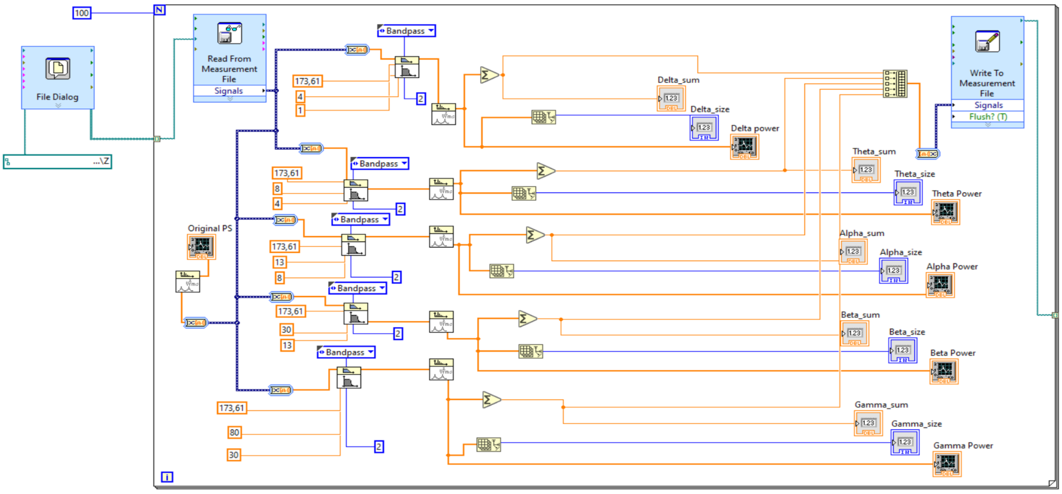


Figure 1: Labview Scheme

B. *Feature Extraction*

In this paper, the total power spectrum of delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) sub-bands of EEG signals were obtained using LabVIEW. At first step file path was determined using "File Dialog", a for loop was created, and the "Read From Measurement File" function was used to read text files. Sub-bands were obtained using appropriate bandpass filters as 1-4 Hz for δ, 4-8 Hz for θ, 8-13 Hz for α, 13-30 Hz for β and 30-80 Hz for γ frequency band [6]. Power spectrums of these 5 sub-bands were obtained, the total value of the spectrum was calculated and written on an excel file. This procedure was performed for both epileptic signal dataset 'E' and dataset 'A' that contains the recording of healthy volunteers. The architecture of the LabVIEW system for feature extraction process is represented in Figure 1.

C. *Classification*

Machine learning algorithms are widely used in classification and regression problems. There exist several methodologies to differentiate a sample from other groups and determine the class of the sample. In this paper Naïve Bayes, Logistic Regression, Quadratic Discriminant, Support Vector Machine (SVM), and K-nearest Neigbors (KNN) classifier algorithms were chosen after comparing their performances with other classifier types.

1) Naive Bayes Classifier: Naïve Bayes classifier is a pattern recognition technique that uses a probabilistic approach and the fundamental acceptance is that all features are independent of each other and assumed to be equally important [7]. In this paper, the Gaussian Naïve Bayes algorithm was used which means it is assumed that all features are continuous and they are sampled from a Gaussian distribution.

2) Logistic Regression: In Logistic Regression algorithm, maximum likelihood is calculated and a sigmoid function is used to classify the sample, thus the output is always a binary value such as 0/1 [8].

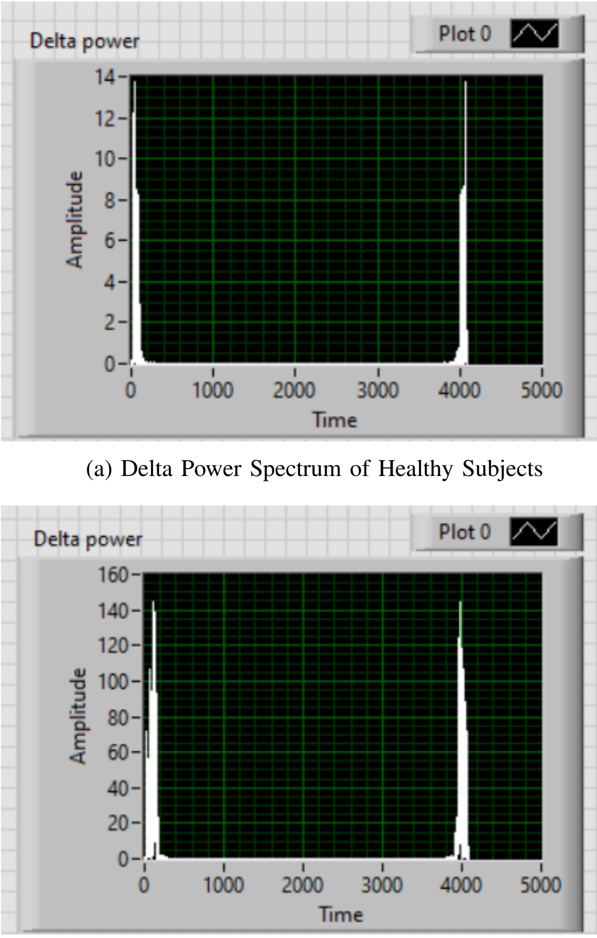
3) Quadratic Discriminant: Quadratic Discriminant analysis is a different version of Linear Discriminant analysis where classification is performed o separate classes with a quadratic surface and nonlinear analysis is adopted for this reason. The assumption is that the classes are normally distributed and covariances of classes are not specifically identical [9].

4) Support Vector Machine: SVM is a machine learning technique based on creating a hyperplane between classes. SVM normally separates data linearly however by using a kernel function, non-linear analysis can also be performed. Since it is a supervised learning algorithm, data is labeled. The closest data points to hyperplane are called 'support vector' and the distance between them is called 'margin'. The best fitting hyperplane is selected by making margin as wide as possible so that a sharp separation between classes can be performed [7]. In this paper, fine Linear SVM was used classification.

5) K-Nearest Neighbors: K-NN algorithm is a pattern recognition technique that adopts the idea that the related samples stay together. It calculates the distance of the samples to K neighbors and the class of the sample is determined as the most related ones [7]. In this study, Euclidian distance and medium KNN was used.

III. Results and Discussion

In this paper, the total power of δ, θ, α, β, γ sub-bands were obtained from power spectrum and used as features. In this way, epileptic and normal EEG signals were differentiated using Naïve Bayes, Logistic Regression, Quadratic Discriminant, SVM, and KNN classifiers. 5-fold cross-validation resampling procedure was used. Sensitivity is the measure of the correctness of the identification of subjects with the disease which is calculated with the formula given in equation 1, specificity is the measure of the correctness of the identification of the subjects without the disease which is calculated with the formula given in equation 2. Accuracy is the ability



(a) Delta Power Spectrum of Healthy Subjects

(b) Delta Power Spectrum of Epileptic Subjects

Figure 2: Delta Power Spectrum of Healthy and Epileptic Subjects

to differentiate both categories and it is calculated with the formula given in equation 3 [10]. The number of related variables is given in Table 1 where TP is True Positive, FP is False Positive, TN is True Negative and FN is False Negative.

$$Sensitivity = \frac{TP}{TP+FN} \tag{1}$$

$$Specificity = \frac{TN}{TN+FP} \tag{2}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

Since accuracy is not a reliable source for the performance of the system alone, in Table 2, along with the accuracy, sensitivity, and specificity values are given for each classifier. As can be seen, the KNN classifier had the highest accuracy with 95.1%. In Figure 2, the δ Powers of both healthy and epileptic subjects are given as an example. It can be seen that

Classifier	TP	TN	FP	FN
Quadratic Discriminant	446	497	3	54
Logistic Regression	449	495	5	51
Naive Bayes	446	497	3	54
SVM	436	500	None	64
KNN	458	493	7	42

Table I: Results of Confusion Matrix

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
Quadratic Discriminant	94.3	89.2	99.4
Logistic Regression	94.4	89.8	99
Naive Bayes	94.3	89.2	99.4
SVM	93.6	87.2	100
KNN	95.1	91.6	98.6

Table II: Results of the Classification

the epileptic δ power was significantly higher than the δ power of healthy subjects.

In a study conducted by Subasi, the same EEG dataset was used and WT was applied on the dataset to obtain sub-bands. Mixture of Experts (ME) model and Multi Layer Perceptron Neural Networks (MLPNNs) were used for the classification of seizure and normal EEG signals. Correct classification performances were 94.5% for ME model and 93.2% for MLPNN model [11]. This indicated that the difference between the correct classification for the study conducted by Subasi and the accuracy of the models used in this paper was negligible. However, it should be noted that it is hard in terms of model complexity to use an ANN model instead of ML classifiers.

IV. Conclusion

In this study, LabVIEW was used to calculate total powers of each EEG sub-bands as δ, θ, α, β and γ from power spectrum to use as features in order to differentiate epileptic and normal EEG signals using classifiers as Logistic Regression, Naïve Bayes, Quadratic Discriminant, SVM and KNN. The maximum accuracy value was obtained using KNN as 95.1%. It was concluded that the total power of each sub-bands can be used for the identification of epileptic patients from EEG signals. The future work can be done in order to determine the dominant sub-band for epileptic seizures.

References

[1] Teplan, M. (2002). Fundamentals of EEG measurement. Measurement science review, 2(2), 1-11.

[2] World Health Organization, Global Campaign against Epilepsy, Programme for Neurological Diseases, Neuroscience (World Health Organization), International Bureau for Epilepsy, World Health Organization. Department of Mental Health, ... & International League against Epilepsy. (2005). Atlas: epilepsy care in the world. World Health Organization.

[3] Acharya, U. R., Sree, S. V., Swapna, G., Martis, R. J., & Suri, J. S. (2013). Automated EEG analysis of epilepsy: a review. Knowledge-Based Systems, 45, 147-165.

[4] Chua, K. C., Chandran, V., Acharya, R., & Lim, C. M. (2008, August). Automatic identification of epilepsy by HOS and power spectrum parameters using EEG signals: A comparative study. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 3824-3827). IEEE.

[5] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. Elger, Phys. Rev. E 64, 061907 2001!.

[6] Tsipouras, M. G. (2019). Spectral information of EEG signals with respect to epilepsy classification. EURASIP Journal on Advances in Signal Processing, 2019(1), 10.

[7] Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, 3-24.

[8] Subasi, A., & Ercelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. Computer methods and programs in biomedicine, 78(2), 87-99.

[9] Tessy, E., Shanir, P. M., & Manafuddin, S. (2016, September). Time domain analysis of epileptic EEG for seizure detection. In 2016 International Conference on Next Generation Intelligent Systems (ICNGIS) (pp. 1-4). IEEE.

[10] Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. NESUG proceedings: health care and life sciences, Baltimore, Maryland, 19, 67.

[11] Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Systems with Applications, 32(4), 1084-1093.