Epileptic Seizure Detection based on EEG Data Using Discrete Wavelet Transform

Hakan BİÇER

Eskişehir Osmangazi University

Department of Electrical and Electronics Engineering

Introduction

- **Epilepsy** is a chronic disorder that causes unprovoked, recurrent **seizures**.
- At least 40-50 million people worldwide (about 1% of population) suffer from Epilepsy.
- For 6 out of 10 people with epilepsy, the cause can't be determined.
- A **seizure** is a sudden rush of electrical activity in the brain.
- **Epileptic seizures** occur when a massive group of neurons in the cerebral cortex suddenly begin to discharge in a highly organized rhythmic pattern.
- Seizures usually happen **spontaneously**, in the absence of external triggers.
- Seizures cause temporary disturbances of brain functions such as **motor control, responsiveness and recall** which typically last from seconds to a few minutes.

Source: "Everything You Need to Know About Epilepsy". https://www.healthline.com/health/epilepsy

How is epilepsy diagnosed?

- **Electroencephalogram (EEG)** is the most common and effective test used in diagnosing epilepsy.
- Electrodes are attached to scalp.
- The electrodes record the electrical activity of your brain.
- Skilled neurophysiologists visually examine the EEG signals and detect epilepsy.
- Epileptic seizures can be detected by analyzing long recordings of EEG signals.

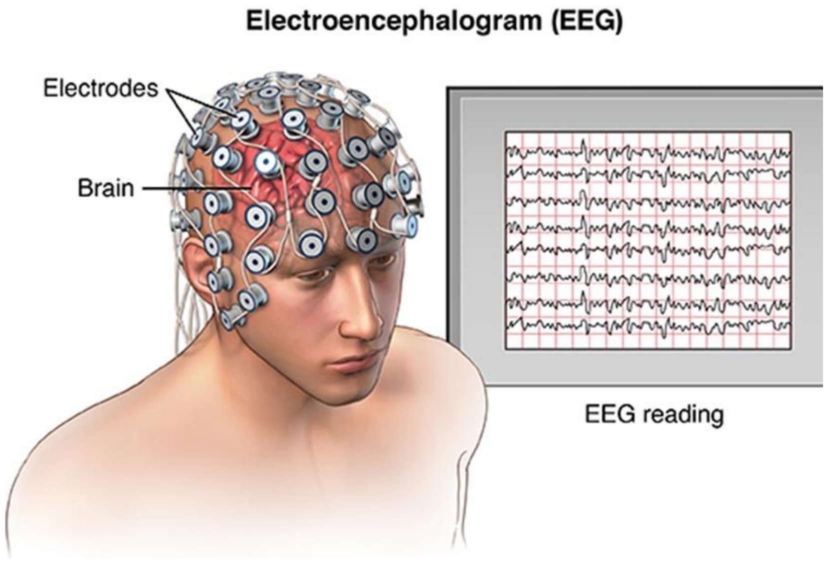

Photo Credit: http://drmridha.com/services/eeg

Main Aim and Structure of This Project

- Main Goal: Classifying raw EEG data as "with-seizure" and "without-seizure.

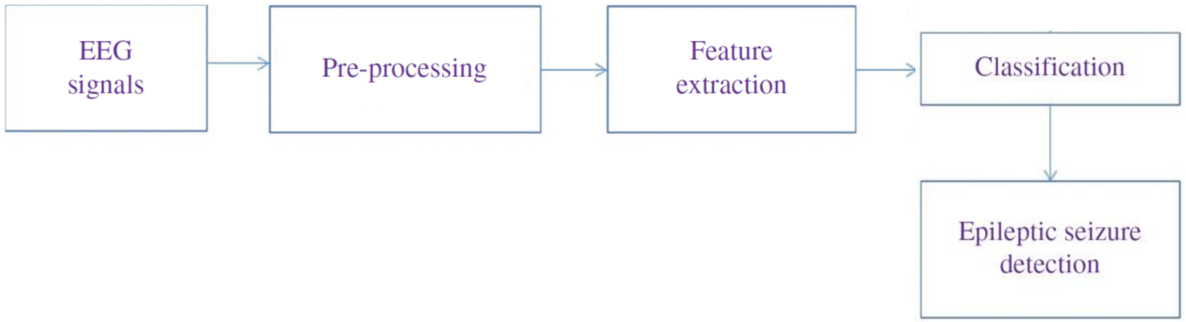

Figure 1. General structure of EEG pattern recognition.

- Pre-processing: Channel reduction, Discrete Wavelet Transform
- Feature Extraction: Max, Min, Mean, etc.
- Classification: SVM and Random Forest

Dataset

- **Publicly available** CHB-MIT Scalp EEG Database collected at the Children's Hospital Boston, Massachusetts (MIT).
- There are **24 cases**, each contains between 9 and 42 **continuous** .edf files from a single subject.
- .mat files can be found in PhysiobankATM belonging to Physionet.
- The EEG signals were sampled at **256Hz**.
- In most cases, the .mat files contain exactly **one hour** of digitized EEG signals (**921.600 samples**).
- Most files contain **23 EEG signals**, which means dimension of 1 .mat file is 23 x 921.600 matrix.
- However, there are some dummy (named “-”), ECG and VNS signals in some files which were excluded in this study.
- Annotations only involve seizure start and end time, and they are comprehensible and clear.
  [231456 250121]

  Seizure start Seizure end

Examples of the Dataset

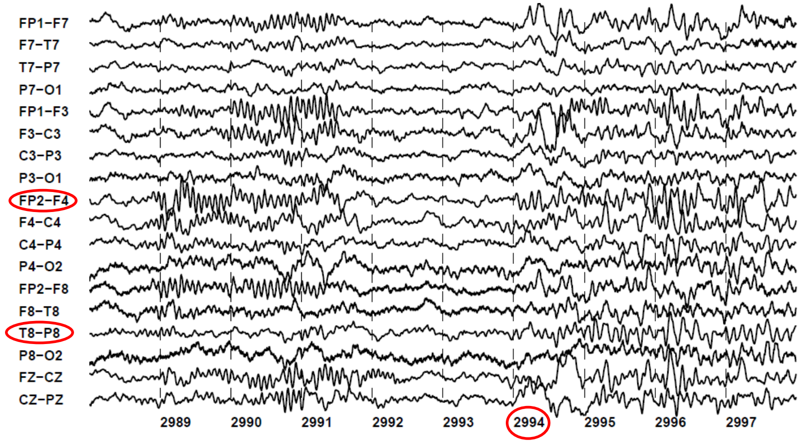

Figure 2. A seizure within the scalp EEG of Patient A (Shoeb, 2009).

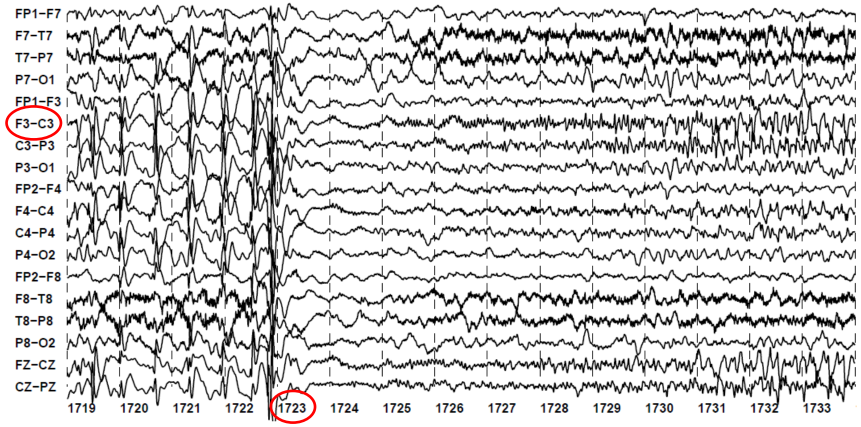

Figure 3. A seizure within the scalp EEG of Patient B (Shoeb, 2009).

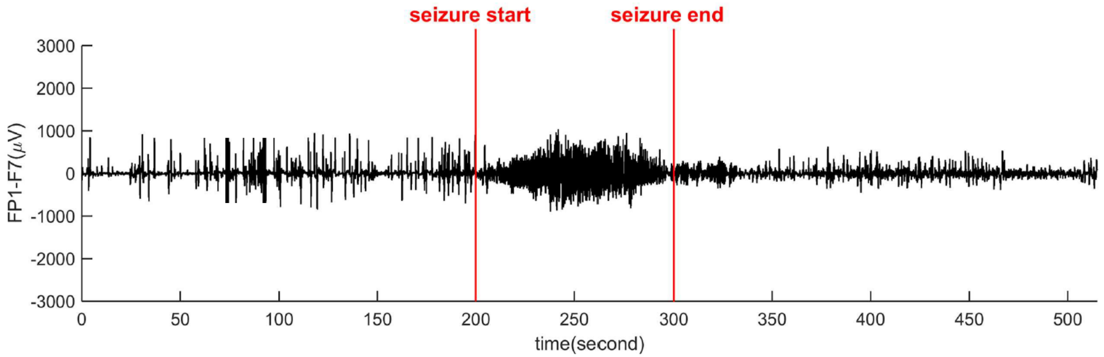

Figure 4. Seizure and non-seizure EEG segments from MIT dataset (Chen et al., 2017).

Source: Shoeb, A.H. (2009). Application of machine learning to epileptic seizure onset detection and treatment.
Chen D, Wan S, Xiang J, Bao FS (2017) A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PLoS ONE 12(3): e0173138.

Data Preparation

- **Seizures more than 20 seconds** were extracted from the files with seizure.
- Data with seizures was split into **2 second EEG epochs** to obtain **stationary data**.
- In order to keep data balanced, 16 and 4 second segments were extracted from every file without seizure for training and testing, respectively.
- They were also split into 2 second epochs.
- Consequently, **10464 EEG epochs** with 23 channels were used from 24 cases in total.

Dimension of 2-second epoch: 23 x 512

Pre-processing

Fourier Transform

- Frequency domain method
- Provides **only spectral information** in frequency domain
- Loses information in time domain

Discrete Wavelet Transform

- Time-frequency domain method
- Provides information about both **frequency** and **location** in time domain with high resolution
- Represents continuous signals **more accurately** and has **less information loss**

Discrete Wavelet Transform

- A **discrete wavelet transform (DWT)** is a transform that decomposes a given signal into a number of sets, where each set is a time series of coefficients describing the time evolution of the signal in the corresponding frequency band.

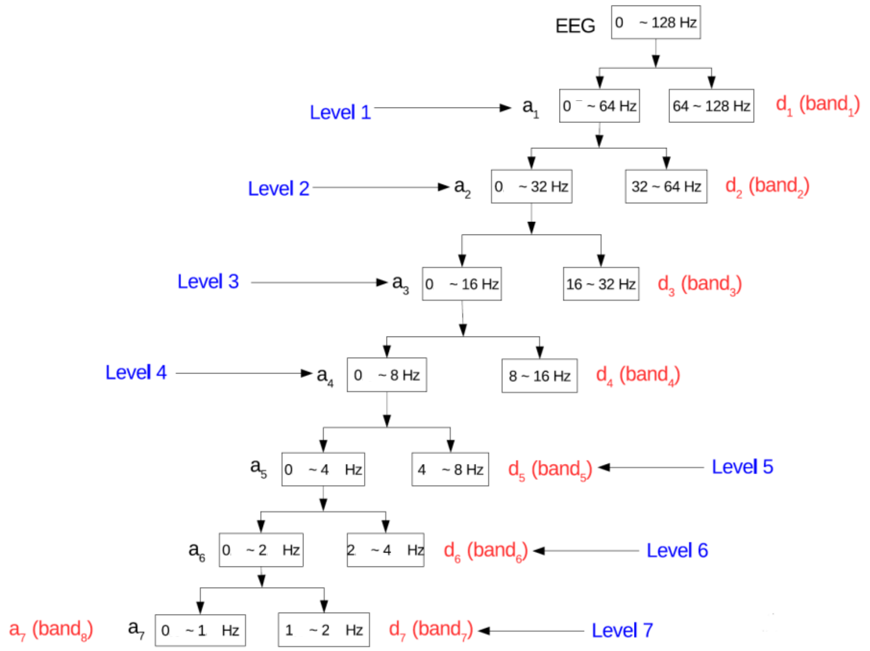

Figure 4. Structure of 7-level wavelet decomposition of CHB-MIT Dataset (Chen et al., 2017).

Wavelet family	Mother wavelet
Biorthogonal (bior)	bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8
Coiflets (coif)	coif1, coif2, coif3, coif4, coif5
Daubechies (db)	db1, db2, db3, db4, db5, db6, db7, db8, db9, db10
Reverse biorthogonal (rbio)	rbio1.1, rbio1.3, rbio1.5, rbio2.2, rbio2.4, rbio2.6, rbio2.8, rbio3.1, rbio3.3, rbio3.7, rbio3.9, rbio4.4, rbio5.5, rbio6.8
Symlets (sym)	sym2, sym3, sym4, sym5, sym6, sym7, sym8
Discrete Meyer (dmey)	dmey
Haar (Haar)	haar

Figure 5. Fifty-four Mother Wavelets (Chen et al., 2017).

Source: Chen D, Wan S, Xiang J, Bao FS (2017) A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PLoS ONE 12(3): e0173138.

Feature Extraction

- The most discriminative features for CHB-MIT dataset: Energy, Max, Min, Mean, Standard Deviation, Normalized Standard Deviation and Skewness (Chen et al., 2017).
- 6 features were extracted for 6 sub-bands.
- 6 x 6 = **36** features for **1 channel**.
- For **23 channel**, 23 x 36 = **828** features.
- Therefore, **each epoch** can be represented a feature vector with **dimension of 1x828.**

Source: Chen D, Wan S, Xiang J, Bao FS (2017) A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PLoS ONE 12(3): e0173138. doi:10.1371/ journal.pone.0173138

Classification

- Feature vectors were classified as “seizure” and “non-seizure”.
- Linear SVM and Random Forest Classifiers were used.
- 8322 feature vectors were trained in total.
- 2142 data were tested on both model.

	Training	Test
With-seizure	4170 epochs	1104 epochs
Without-seizure	4152 epochs	1038 epochs

Results

	SVM	RANDOM FOREST
Sensitivity	92.00%	97.74%
Specificity	79.98%	93.72%
Positive Predictive Value	78.17%	93.84%
Negative Predictive Value	92.78%	97.69%
False Positive Rate	20.02%	6.29%
False Negative Rate	8.00%	2.26%
False Discovery Rate	21.83%	6.16%
False Omission Rate	7.23%	2.31%
Accuracy	85.25%	95.71%

- **Random Forest** classifier outperformed **SVM** with **95.71%** and **85.25%** of accuracy, respectively.

Big O Notation

- **The complexity of SVM:**
  - Training Time Complexity = $O(n^2)$ n = number of training examples
    - n = 8322
    - $O(8322^2)$

- **The complexity of Random Forest:**
  - Training Time Complexity = O(n * log(n) * d * k) d = dimentionality of the data, k = number of Decision Trees
    - n = 8322
    - d = 828
    - k = 20
    - O(8322 * log(8322) * 828 * 20)

Conclusion

- Detecting epileptic seizures manually from very long scalp EEG recordings is **demanding** and **challenging** for doctors.
- In this work, **DWT based method** was proven that it is possible to detect seizures **automatically** and **effectively**.
- Automatic seizure detection can be done successfully with more than 85% of accuracy depending on the classifier.
- This work might help doctors to discriminate and detect seizures **in a very short time**.
- Undoubtedly, it should be improved in further studies in terms of **speed, computational complexity** and **accuracy**.

THANK YOU !