Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230426

Multi-Class Seizure Type Classification Using Features Extracted from the EEG

Abirami SELVARAJ[a,1], Swarubini PJ[b], John THOMAS[c], Yuvaraj RAJAMANICKAM[d], Ramshekhar N MENON[e], and Jac Fredo AGASTINOSE RONICKOM[a]

[a] *School of Biomedical Engineering, Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh, India*

[b] *Department of Sensor and Biomedical Technology, VIT University, Vellore campus, Vellore, India*

[c] *Montreal Neurological Institute, McGill University, Canada*

[d] *Science of Learning in Education Centre (SoLEC), National Institute of Education (NIE), Nanyang Technological University (NTU), Singapore*

[e] *Department of Neurology, R Madhavan Nayar Centre for Comprehensive Epilepsy Care, Cognition Behavioral Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India*

ORCiD ID: Jac Fredo AGASTINOSE RONICKOM https://orcid.org/0000-0001-5759-6632

Abstract. In this study, we classify the seizure types using feature extraction and machine learning algorithms. Initially, we pre-processed the electroencephalogram (EEG) of focal non-specific seizure (FNSZ), generalized seizure (GNSZ), tonic-clonic seizure (TCSZ), complex partial seizure (CPSZ) and absence seizure (ABSZ). Further, 21 features from time (9) and frequency (12) domain were computed from the EEG signals of different seizure types. XGBoost classifier model was built for individual domain features and combination of time and frequency features and validated the results using 10-fold cross-validation. Our results revealed that the classifier model with combination of time and frequency features performed well followed by the time and frequency domain features. We obtained a highest multi-class accuracy of 79.72% for the classification of five types of seizure while using all the 21 features. The band power between 11-13 Hz was found to be the top feature in our study. The proposed studycan be used for the seizure type classification in clinical applications.

Keywords. Epilepsy, multiclass seizure, feature extraction, XGBoost classifier

1. Introduction

Seizure is an uncontrolled, rapid electrical disturbance in the brain which makes the

[1] Corresponding Author: Abirami Selvaraj, email: abiramis.bme20@itbhu.ac.in

patient vulnerable and risky [1]. It is medically divided into two primary groups depending on the degree to which brain areas are impacted as focal and generalized seizures. Focalseizures are subdivided into simple and complex seizures based on the patient's level of awareness, affecting a specific area of the brain. Generalized seizures that affect the majority of the brain are further classified as absence, tonic, atonic, clonic, tonic-clonic, or myoclonic based on motor and non-motor symptoms [2]. Classification of seizure is very essential for accurate diagnosis and treatment. Electroencephalogram (EEG) is the most practical and cost-effective tool to diagnose epilepsy among the presently available brain imaging techniques. The factors such as variability of symptoms across patients, poor inter-rater agreement, signal artefacts add to the hindrance of the accurate interpretation of the EEG recording [3]. The manual mapping of the seizure type data with the help of clinical information involves huge amounts of time, resources and is prone to human errors. Hence, an automated seizure type classification model is essential for im- proving the effectiveness of the seizure patient management. In this study, we attempted to classify the five seizure types such as focal non-specific seizure (FNSZ), generalized seizure (GNSZ), tonic-clonic seizure (TCSZ), complex partial seizure (CPSZ) and ab- sence seizure (ABSZ) using the feature extraction and machine learning models.

2. Methods

The process pipeline followed in this study is depicted in 1.

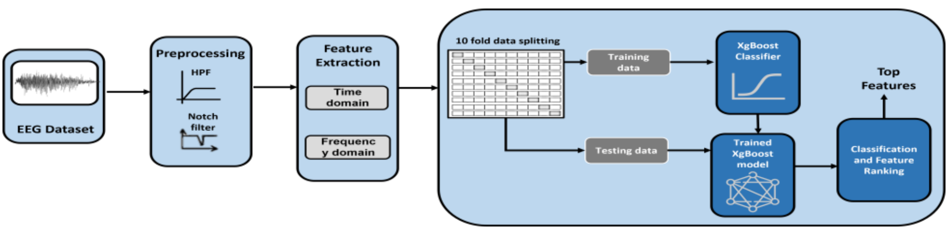

Figure 1. Pipeline of the study

We considered the EEG signals publicly available in the temple university seizure corpus (TUSZ) ver. 1.5.2 dataset for our analysis [4]. The dataset consists of 3,050 seizures with diverse morphologies recorded from 300 patients. The recordings are of different seizure duration (mean: 97 seconds) and sampling rate (250-1000Hz).

Table 1.: List of features

Feature Domain	Features
Time domain Features	Root mean square, standard deviation, Hjorth mobility, teager energy, Higuchi fractal, hurst exponent, mean line length, zero crossing rate, entropy
Frequency domain Features	Delta band power, high gamma band power, band power 0-2 Hz, band power 2-4 Hz, band power 14-16 Hz, band power 19-21 Hz, band power 7-9 Hz, band power 10-12 Hz, band power 11-13 Hz, band power 8-10 Hz,band power 20-22 Hz,peak frequency

Initially, the EEG segments that are solely responsible for seizures were extracted from the dataset. The start and stop time provided in the annotation file was used for extracting the different types of seizure. In order to preserve all the information in the signal, down sampling was not performed. Bipolar montage was applied to get the difference of the electrode potentials. We preprocessed the considered signals using a notch filter of 60 Hz and a high pass filter (HPF) of 1 Hz [5]. Further, 9 time and 12 frequency domain features given in Table 1 were extracted from the seizure segments. The XG-Boost classifier was employed for the classification of the five types of seizures such as FNSZ, GNSZ, CPSZ, TCSZ and ABSZ. The performance of the classification model was cross-validated using a 10-fold method to evaluate the multi-class model accuracy (True positives / Total). We analyzed the effect of the classification model on time (9), frequency (12) and combination of time and frequency (9+12) features and identified the most important feature contributing to the classifier. The feature importance score of each feature was identified by averaging feature importance of each feature across 10-folds.

3. Results and Discussions

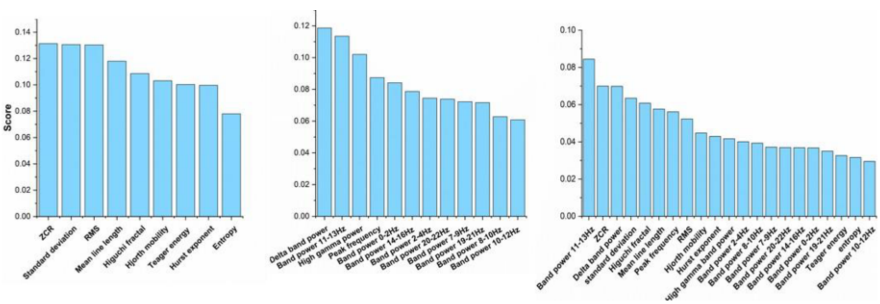

Figure 2. Feature ranking result of XGBoost on time (left), frequency (middle), and all features (right)

The feature ranking results using XGBoost feature importance method for time, frequency domain and time+frequency features are shown in Figure 2 (a-c). It can be observed that the zero-crossing rate, standard deviation, and root mean square were the top three-time domain features obtained by XGBoost ranking. Further, frequency domain features such as delta band power, band power 11-13 Hz and high gamma band power contributed well for the classification model. While we supplied all the time and frequency domain features to the XGBoost, we received the top five features as band power 11-13 Hz, zero crossing rate, delta band power, standard deviation, and Higuchi fractal. It can be noted that most of the top 5 features were from the time domain although the top feature is from the frequency domain. Our results show the influence of time and frequency domain features in the multi-class seizure classification. The classification accuracy was high while feeding the combination of time and frequency features to the model followed by the time and frequency domain features. We achieved a maximum classification accuracy of 79.72%. However, the average 10-fold cross-validation accuracy was 65.88% for the classification of five types of seizure using time+frequency features. Our study highlights the importance of utilising the combination of time and frequency features from time and frequency domain for the classification model.

This study can be extended by including more time, frequency and time-frequency features that would help in improving the performance of the model. A comparative study on different machine learning classifiers can be implemented to identify the best type of classifier for seizure type classification. After the examination of the best features, the model accuracy can be further improved by using deep learning neural techniques. The results of this study highlight the importance of the use of combination of different domain features, this can be further extended for the identification of the best combination of features. The same study can be optimized further by comparing the performance with different scalp EEG seizure type dataset.

4. Conclusions

In this study, we investigated the performance of time and frequency domain features of EEG for multi-class seizure type classification. The XGBoost gradient boosting algorithm was implemented to evaluate the model performance and identify the important features. Our results indicate that the use of a combination of features from time and frequency domain increases the performance of the model. It can be inferred from our study that the use of best distinguishing features can fetch higher results. We obtained a highest of 79.72% classification accuracy and an average 10-fold accuracy of 65.88% with the use of a minimal number of features. Our results show that the band power 11-13 Hz significantly contributes to the classifier. The proposed model with the combination of different domain features has potential application value for clinical systems.

References

[1] Thijs, R.D. et al. (2019) 'Epilepsy in adults', The Lancet, 393(10172), pp. 689–701. Available at: https://doi.org/10.1016/S0140-6736(18)32596-0Fisher.

[2] RS, Cross JH, French JA, Higurashi N, Hirsch E, Jansen FE, Lagae L, Moshe´ SL, Peltola J, Roulet Perez E, Scheffer IE. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017 Apr;58(4):522-30.

[3] Ahmedt-Aristizabal D, Fernando T, Denman S, Petersson L, Aburn MJ, Fookes C. Neural memory networks for seizure type classification. In2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) 2020 Jul 20 (pp. 569-575). IEEE.

[4] Obeid I, Picone J. The temple university hospital EEG data corpus. Frontiers in neuroscience. 2016 May 13;10:196.

[5] Abirami S, Thomas J, Yuvaraj R, Agastinose Ronickom JF. A Comparative Study on EEG Features for Neonatal Seizure Detection. InBiomedical Signals Based Computer-Aided Diagnosis for Neurological Disorders 2022 Jun 18 (pp. 43-64). Cham: Springer International Publishing.