 DOI: 10.3217/978-3-85125-962-9-9

EEG Channel Selection Based on Feature Importance for Epileptic Seizure Classification

Marie Øverby[1*], Luis Alfredo Moctezuma[2], Marta Molinas[1]
[1] *Norwegian University of Science and Technology, Trondheim, Norway*
[2] *University of Tsukuba, Tsukuba, Ibaraki, Japan*
* O. S. Bragstads Plass 7034, Trondheim, Norway. E-mail: marieov@stud.ntnu.no

Introduction: Classification of epileptic seizures based on electroencephalography (EEG) is a well-established research area. Most research uses a patient-dependent approach, i.e., training and test data come from the same patient. In order to use the results on other patients, a patient-independent model needs to be developed. The patient-independent model has to combine high accuracy, few features, and few channels in order to be fast and easy to use. This work presents both a patient-dependent and a patient-independent approach for epileptic seizure detection.

Material, Methods and Results: CHB-MIT and Siena Scalp EEG databases were used [1, 2]. Both use the 10-20 system for electrode placement. The EEG signals from the CHB-MIT database were first decomposed into four sub-bands using the Discrete Wavelet Transform (DWT), from the sub-bands sixteen features were extracted. To classify seizure and seizure-free periods, the features were used as input for Random forest (RF), Gradient Boosting (GB), and Support Vector Machine (SVM). The patient-dependent method was used as a starting point for the patient-independent method and to verify the high performance. Verification included balancing and unbalancing the dataset, changing the size of the training set, using different performance measures, and running the algorithm on the Siena Scalp EEG Database.

For the patient-independent method, the feature importance for each machine learning method was found using Mean Decrease in Accuracy (MDA), and the most important features were chosen. These features were used to find the channel importance using Mean Decrease in Impurity (MDI). Petrosian fractal dimension was found to be the most important for RG and GB over all channels, while standard deviation, root-mean-square, Katz fractal dimension were found to be the most important for SVM over all channels. The accuracy when using these features and the one most important channel are presented in Table 1. The average accuracy for the patient-dependent method is shown in the same table.

	RF	GB	SVM
Patient-dependent	**0.999**	0.996	0.992
Patient-independent	**0.976**	0.884	0.964
	FT9	F7	FT10

Table 1: Accuracy for epileptic seizure classification. The indicated channel is the one found to be most important. Standard deviation ±3%.

Discussion: RF is the method that gives the best results, although they all perform well even with only one feature and one channel. A clear conclusion on Which feature and channel give the best performance is not given due to variations in results for each run of the method. Results presented are from one run. The analysis of MDI and MDA for the selection of features and channels in different runs, in addition to why different runs yield different performance, is left for further work. Nevertheless, high-performance measures for the patient-dependent and patient-independent method were obtained.

Significance: Using only the most important electrodes, the vision is that patients can be notified on a portable device when a seizure occurs and, in the future, before it occurs.

References:

[1] Amaral L. Glass L. Hausdorff J. Ivanov P.C. Mark R. Mietus J.E. Moody G.B. Peng C.K. Goldberger, A. and H.E. Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.

[2] Paolo Detti. Siena scalp eeg database (version 1.0.0). *PhysioNet*, 2020.