Fully and robust unsupervised detection of epileptic seizure from long-term EEG recordings

Jieun Kim[1], Dong-Uk Hwang[1], Edwin J. Son[1], Whansun Kim[1], and Sang Hoon Oh[1]

1) *Division of Basic Researches for Industrial Mathematics, National Institute for Mathematical Sciences, Daejeon 34047, KOREA*

Corresponding Author: Dong Uk Hwang, duhwang@nims.re.kr

ABSTRACT

According to the World Health Organization (WHO), epilepsy is a neurological disease that affects about 6.5 million people, and it affects 3.52 out of every 1,000 people in Korea. About 30% of them are difficult to treat with drugs, so surgery must be treated. For this, it is important to accurately estimate the location of the epileptic lesion. Electroencephalography (EEG) plays an important role in monitoring brain activity and diagnosing epilepsy in patients with epilepsy, but detecting epilepsy activity requires a specialist analyzing long-term measured EEG records. However, this method is time consuming and requires a tedious process, and it is expensive to train experts. Therefore, we developed an algorithm aiming at fully automatic detection of seizures in long-term EEG data using the K-means clustering method, an unsupervised learning method. The automatic detection algorithm developed so far requires training data, that is, data labeled by an expert. Although detection is performed by unsupervised learning, there are cases where the false positive rate is very bad. In threshold-based analysis, there is a trade-off between sensitivity and false positives rate. Other studies do not have an automatic artifact removal method. In order to solve these problems, we developed an algorithm for automatically detecting and removing artifacts in the preprocessing stage, extracting features of ictal waves that appear during seizures, applying k-means classification, and extracting events continuously detected in the post-processing stage. As a characteristic of the ictal wave, the average power spectral density (PSD) of two frequency bands of 1 to 5 Hz and 6 to 13 Hz was used. We have defined three artifacts that interfere with finding the seizure in EEG; Power line noise, slow sleep wave, inter-ictal spikes. Artifacts were detected and removed using the features of each artifact. The ictal wave signal and the background signal were classified using curvature from clear data scattered in 2D. To solve the unbalanced data, oversampling was used and then k-mean clustering was performed. In the post-processing step, only events detected continuously were filtered. In this process, we applied an adjusted threshold and obtained a sensitivity of 84.9% and a 0.093 false positive rate/hour for 21 subjects in the CHB-MIT data set. Our study showed superior results compared to previous actual unsupervised seizure detection studies.