V.O. Gaidar, O.O. Sudakov

Taras Schevchenko National University of Kyiv
E-mail: gaidar.viktoria@gmail.com, sudakov.oleksandr@gmail.com

Statistical sampling and feature selection for epilepsy pattern recognition

Epilepsy is one of the most common neurological diseases that has broad spectrum of debilitating medical and social consequences. The automatic forecasting and detecting systems are vitally important, since they allow patients to avoid dangerous activities in advance of the seizure. We present some methods of feature extraction and selection for detecting the epileptiform activity in electroencephalography signals, based on the processing of a non-stationary signal. The proposed approach is based on the application of the Discrete Wavelet Transform (DWT) and signal processing techniques in order to create the feature vector. Afterwards, the principal component analysis and support vector machine techniques are used in order to reduce the dimensionality of the feature vector.

Keywords: *electroencephalogram, wavelet transform, epileptiform pattern, feature ranking.*

The heaviest burden of epilepsy is that people with this disease are not aware when the next seizure will come. Electroencephalography (EEG) is one of the most useful techniques in order to detect abnormalities, which attends of the epilepsy seizure. It is a safe and multifunction medical tool that is used to localize and identify the epilepsy syndrome [1]. EEG measures the electrical activity of brain.

Within the goal of seizure detection and prediction by the apportion of EEG signals into the ictal (seizure) and pre-ictal (before seizure) periods, we need to solve two major tasks. One of them is the feature extraction from EEG signal that provides reliable data for discriminative analysis. The second one is to determine the approach that will be used to solve the classification problem with high accuracy. Here, we concentrate on the first task.

Materials and methods. The datasets presented in this article are the time series of one-channel intracranial recordings collected from Bogomoletz Institute of Physiology in Kyiv, Ukraine. The data are acquired at a sampling rate of 416 Hz using gold plated electrodes. The data were recorded from fifteen laboratory rats with epilepsy, in average from 50 min to 2.5 hours for
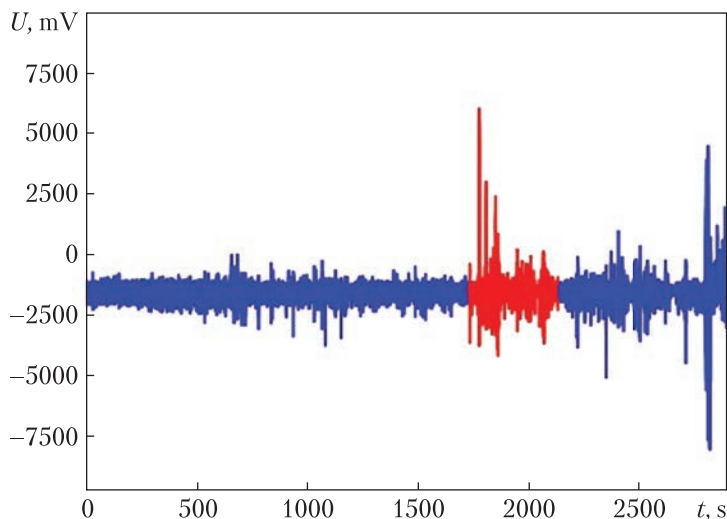
Fig. 1. EEG recording with seizured and non-seizured periods

one examination. The signals were filtered between 0.5 to 40 Hz and then segmented into pre-ictal and ictal states (Fig. 1).

We divided each EEG signal into epoches (time laps), where the non-seizure and seizure activities occurred and then shuffle them in one database. Each signal contains 10575 datapoints, so one epoch contains 1024 points in average. Next, we extracted fifteen different features from each allocated epoch and applied the proposed feature selection algorithm for selecting the least number of features.

Feature extraction. First of all, the statistical features such as the minimum value, maximum value, mean, variance, skewness, and kurtosis were calculated. Gathered data were conducted as the stochastic variable $X = (x_1, x_2, \dots x_n)$.

Mean value is the average value on a specific time period of EEG signal. Dispersion is calculated as var $(x_1, x_2, \dots x_n)$. The variance measures how far the signal is dispersed from the mean value. The standard deviation shows how far the signal fluctuates from the mean. It quantifies the quantity of variation of a set of values and is defined as the square root of the dispersion.
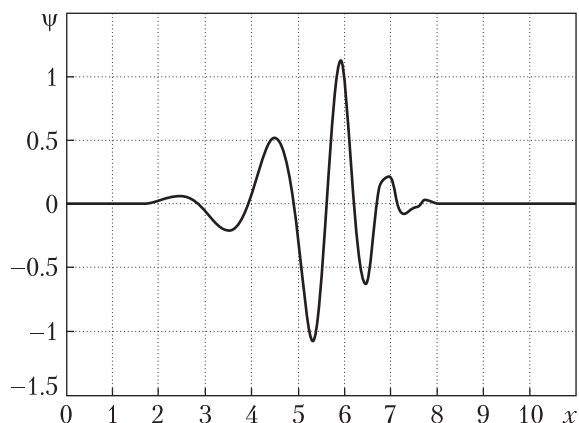
Skewness is a measure of the deviation of a sample distribution from the normal distribution. Skewness determines the symmetry of the probability density function (PDF) of the amplitude of a time series. The height and sharpness of the central peak relative to the rest of the data is measured by kurtosis.

Hjorth parameters are the indicators of statistical properties of time-series. There are three Hjorth parameters: activity, mobility, and complexity. Activity parameter can indicate the surface of the power spectrum in a frequency do-
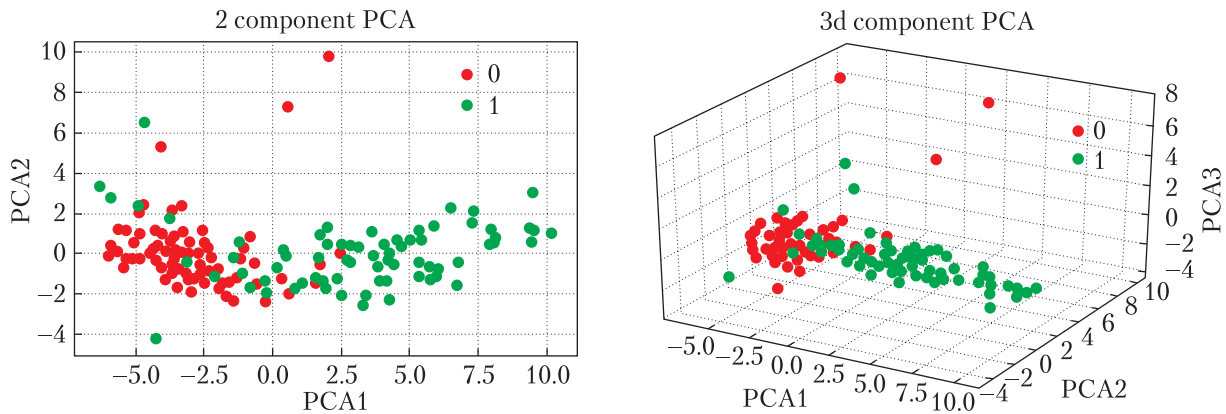

Fig. 2. The mother wavelet function (the Daubechies 5)

54

ISSN 1025-6415. Dopov. Nac. akad. nauk Ukr. 2020. № 4

Fig. 3. The plot of first principal component vectors of EEG features

main. The value of activity returns a large value, if the high-frequency components of the signal exist. It characterizes the mean power of the signal. The mobility parameter represents the mean frequency. Complexity parameter indicates how the shape of a signal is similar to a pure sine wave and represents a change in the frequency. For a pure sine wave, it is zero.

Discrete wavelet analysis. In a normal EEG signal, most of the power is contained within the frequency band from 0 Hz up to 40 Hz. Different pathological and physiological processes are reflected by the activity in different frequency ranges. Discrete Wavelet Transform (DWT) decomposes a signal, as does a Fourier Transform, but in a way that is able to reflect both frequency and temporal location properties of the signal [3].

The DWT is a decomposition of the time series. The method is computationally fast and can be implemented by successive filter banks and thus is an important tool in the signal and image processing. In context of EEG analysis, two wavelets are commonly used: the Daubechies wavelet (Fig. 2) and Symlet wavelet [4].

Results. The principal component analysis has been implied in order to classify the EEG epoch in two different classes — seizure and non-seizured. The PCA analysis has shown that only five features can be used to provide a robust binary classification. Among them are complexity, kurtosis, mean values of approximate wavelet coefficients, energy, and entropy. The results for a 3d plot of PCA vectors are shown in Fig. 3.

For the assessment of a performed algorithm, we compared the results with those of the convensional correlation-based feature selection algorithm. The first detectable discharges often occur with frequencies of about 15 Hz. This discharge frequency decreases during the ictal period, whereas the amplitude increases. The decrement of the discharge frequency is crucial to identify these patterns as epileptic. A minimum ictal phase of five seconds and a minimum inter-ictal phase of ten seconds are assumed. We observed that the proposed algorithm provides a classification accuracy of 77.6 % and sensitivity of 85.7 %. Those results show a similar accuracy with a smaller number of features. Our results indicate that seizure events in a rat model can be detected with high accuracy using the stationary wavelet analysis. The true positive rate of 78 % and true negative rate of 86 % of the proposed algorithm are commonly considered as good enough for different practical applications. The proposed algorithm may be applied to a further development of seizure prediction techniques.

ISSN 1025-6415. Допов. Нац. акад. наук Укр. 2020. № 4

55

REFERENCES

1. Federico, P., Abbott, D. F., Briellmann, R. S., Harvey, A. S. (2005). Functional MRI of the pre-ictal state. Brain, No. 128, pp. 1811-1817.
2. Mormann, F. (2005). On the predictability of epileptic seizures. Clinical Neurophysiology, No. 116(3), pp. 569-587.
3. Suárez, C. C. B. (2010). Wavelet transform and cross-correlation as tools for seizure prediction. Engineering in Medicine and Biology Society (EMBC), International Conference of the IEEE. – IEEE– pp. 4020-4023.
4. Schiller S., Schiller, U., Heisig, K. Goedicke (1977). Use of the ringgap plasmatron for high rate sputtering. Thin Solid Films., No. 4, pp. 327-334.

Гайдар В.О., Судаков О.О.

Київський національний університет ім. Тараса Шевченка

E-mail: gaidar.viktoria@gmail.com, sudakov.oleksandr@gmail.com

ВИДІЛЕННЯ ХАРАКТЕРИСТИЧНИХ ПАРАМЕТРІВ ЕЛЕКТРОЕНЦЕФАЛОГРАФІЧНОГО СИГНАЛУ ДЛЯ ЗАДАЧІ БІНАРНОЇ КЛАСИФІКАЦІЇ ЕПІЛЕПТИЧНИХ СТАНІВ

Епілепсія — одне з найпоширеніших неврологічних захворювань, яке має широкий спектр когнітивних та соціальних проявів. У представленій статті застосовано методи виділення та відбору статистичних ознак інтракраніального електроенцефалографічного сигналу для виявлення епілептичної активності та її ранньої діагностики. Запропонований підхід базується на застосуванні дискретного вейвлет-перетворення та методах обробки сигналів для створення вектора ознак. Метод головних компонент та метод опорних векторів використано для зменшення розмірності вектора ознак та бінарної класифікації. За результатами роботи вектор з тринадцяти компонент було зменшено до вектора з п'яти компонент зі збереженням специфічності та чутливості класифікації.

Ключові слова: *електроенцефалограма, вейвлет-перетворення, епілептичні патерни, вектор ознак.*

Гайдар В.О, Судаков О.О.

Киевский национальный университет им. Тараса Шевченка

E-mail: gaidar.viktoria@gmail.com, sudakov.oleksandr@gmail.com

ВЫДЕЛЕНИЕ ХАРАКТЕРИСТИЧЕСКИХ ПАРАМЕТРОВ ЭЛЕКТРОЭНЦЕФАЛОГРАФИЧЕСКОГО СИГНАЛА ДЛЯ ЗАДАЧИ БИНАРНОЙ КЛАССИФИКАЦИИ ЭПИЛЕПТИЧЕСКИХ СОСТОЯНИЙ

Эпилепсия — одно из самых распространенных неврологических заболеваний, которое имеет широкий спектр когнитивных и социальных проявлений. В представленной статье применены методы выделения и отбора статистических признаков интракраниального электроэнцефалографического сигнала для обнаружения эпилептической активности и ее ранней диагностики. Предложенный подход основан на применении дискретного вейвлет-преобразования и методах обработки сигналов для создания вектора признаков. Метод главных компонент и метод опорных векторов применены для уменьшения размерности вектора признаков и бинарной классификации. По результатам работы вектор из тринадцати компонент был уменьшен до вектора из пяти компонент с сохранением специфичности и чувствительности классификации

Ключевые слова: *электроэнцефалограмма, вейвлет-преобразования, эпилептические паттерны, вектор признаков*