Machine Learning based EEG Signal Classification

Anjali Kadwe
Department of Electronics & Communication Engineering
Sapkal College of Engineering Nashik, India

Abstract

Epilepsy is a neurological disorder which is characterized by transient and unexpected electrical disturbance of the brain. The electroencephalogram (EEG) is a commonly used signal for detection of epileptic seizures. The proposed method is based on the classification of EEG signal with the less number of sample and more accurately by using the Matlab software. The Bonn University Data set use in this project provide classification of EEG signal by using the latest transform method. The project consist of Extraction of the data from text file, Frequency domain low pass filtering And Feature extraction by three most recent transform such as Coiflet Transform, Stationary Wavelet Transform (SWT) and Walsh Hadamard Transform (WHT). This transformed signal is the classified by KNN ensemble classification. This project provide an overall classification accuracy of 99%.
Keywords: Epilepsy, Extraction of Data from Text File, Frequency Domain Low Pass Filtering, Feature Extraction, Classification of EEG signal

I. INTRODUCTION

Epilepsy is a common brain disorder that, according to an estimate of the World Health Organization, affects almost 60 million people around the world. Approximately one in every 100 persons will experience a seizure at some time in their life. Epilepsy is characterized by the recurrent and sudden incidence of epileptic seizures which can lead to dangerous and possibly life-threatening situations. The seizures are the result of a transient and unexpected electrical disturbance of the brain and excessive neuronal discharge that is evident in the electroencephalogram (EEG) signal representative of the electrical activity of the brain. Consequently, the EEG signal has been the most utilized signal in clinical assessments of the state of the brain and detection of epileptic seizures, and is very important for a proper diagnosis of epilepsy. The detection of epileptic seizures by visual scanning of a patient's EEG data usually collected over a few days is a tedious and time-consuming process. In addition, it requires an expert to analyze the entire length of the EEG recordings, in order to detect epileptic activity.

The major challenge in the classification of MI EEG signals arises due to the fact that the brain signals that are recorded are very small in amplitude. Therefore, events such as eye blink, eye movement, muscular movements, teeth grinding and heart rhythm interfere with the EEG signal resulting in a signal having low signal to noise ratio (SNR). This prevents the decoding system to correctly decode the user thoughts. Various techniques have been proposed by the scientific community aiming to improve the temporal filtering methods, spatial filtering, feature extraction and feature selection techniques, dimensionality reduction techniques and classification algorithms. Several feature extraction techniques such as power spectral density (PSD), common spatial pattern (CSP), statistical features, self-organizing maps (SOM), correlation, spectral coherence and information entropy have been studied. Classifiers such as support vector machine (SVM), k-nearest neighbors (KNN) , random forest (RF), etc. have been explored for classification of MI-EEG signals.

In this paper we are classify the EEG signal with the less number of samples and with more accuracy by using the latest feature extraction method to an epileptic and non-epileptic EEG signal.

II. RELATED WORK

Many automated EEG signal classification and seizure detection systems, using different approaches, have emerged in recent years. Among such studies, Gotman presented a computerized system for detecting a variety of seizures, while Qu and Gotman proposed the use of the nearest-neighbor classifier on EEG features extracted in both time and frequency domains to detect the onset of epileptic seizures. Gigola et al. applied a method based on the evolution of accumulated energy using wavelet analysis for the prediction of epileptic seizure onset from intracranial epileptic EEG recordings, while Adeli et al., Guler et al. and Ubeyli et al.discussed the potential of nonlinear time series analysis in seizure detection. Artificial neural network-based detection systems for diagnosis of epilepsy have been proposed by several researchers. The method put forward by Weng and Khorasani uses the features proposed by Gotman and Wang, namely, average EEG amplitude, average EEG duration, variation coefficient, dominant frequency and average power spectrum, as inputs to an adaptive structured neural network. The method proposed by Pradhan et al. uses a raw EEG signal as an input to a learning vector quantization network. Nigam and Graupe proposed a new neural network model called LAMSTAR (large memory storage and retrieval) network and two time-domain attributes of EEG; namely, relative spike amplitude and spike rhythmicity have been used as inputs for the purpose of detecting seizures. The algorithm proposed by

Kiymiketal. uses a back propagation neural network with periodogram and auto regressive (AR) features as inputs for automated detection of epileptic seizures. Ghosh Dastidar et al. discussed a classification methodology based on wavelet analysis and both radial basis function and Levenberg-Marquardt back propagation neural network. Srinivasan et al. presented an algorithm based on approximate entropy as an input to an artificial neural network classifier, while Subasi used wavelet analysis and mixture of experts, in addition to the artificial neural network, to classify EEG signals and detect seizures.

Due to quite a low understanding of the mechanisms underlying the problem, most existing methods suffer from low accuracy, a high rate of false alarms and missed detections. In addition, due to a lack of reliable standardized data, most reported EEG analysis-based algorithms are performed on a small number of datasets, which often demonstrate good accuracy for selected EEG segments but are not robust enough to adjust to EEG variations commonly encountered in a hospital setting.

In this research, however, a larger number of EEG data sets, which belong to two subject groups, were used: a) healthy subjects (normal EEG), b) epileptic subjects during a seizure-free interval (interictal EEG). The EEG signal classification and seizure detection problem was modeled as a three-group classification problem that could be of great clinical significance. An automated system able to accurately differentiate between normal and interictal EEG signals can be used to diagnose epilepsy, while a system that can accurately differentiate between interictal and ictal EEG signals can be used to detect seizures in a clinical setting. Therefore, the classification algorithm must be able to classify all two groups accurately and at the same time be robust with respect to EEG signal variations across various mental states and subjects. The improvement of the classification accuracy is mainly based on the design of both an appropriate feature space, by identifying combinations of all extracted features that increase the inter-class separation, and classifiers that can accurately classify all two groups of EEG signals based on the selected and reduced feature space. Real EEG recordings were applied to test algorithm performance and the results indicated that the algorithm has a potential to be applied within an automatic epilepsy diagnosis system.

III. METHODOLOGY

A. *Extraction of the Data from Text File:*

In this project we have used the A and E dataset from the Bonn University for detection and classification of the EEG signal and all this project is carried out with the help of MATLAB software so extraction of the data from the text file is the first step in the process as all the flow work of the process is shown in the fig 1 as follows:

B. *Frequency Domain Low Pass Filtering:*

The frequency domain processing techniques are based on modifying the Fourier transform of an imagee format. The basic idea in using this technique is to enhance the image by manipulating the transform coefficient of the image, such as Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), and Discrete Cosine Transform (DCT). This methods advantages includes low complexity of computations, ease of viewing and manipulating the frequency composition of the image and the easy applicability of special transformed domain properties.

An ideal low pass filter deals with the removal of all high frequency values of the Fourier transform that are at a distance greater than a specified distance from the origin of the transformed image. The filter transfer function for the Ideal low-pass filter is given by:

$$H(u,v) = 1 \quad \text{if } D(u,v) < D0$$
$$H(u,v) = 0 \quad \text{if } D(u,v) > D0$$

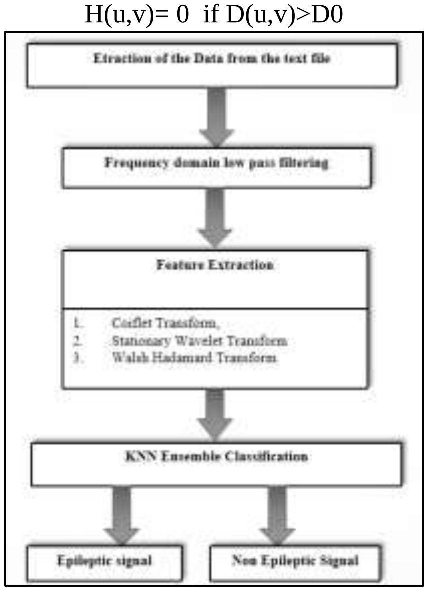


Fig. 1: The work flow of the proposed method.

C. *Feature Extraction*

Feature extraction is either time domain or frequency domain for EEG signal. In most of the cases we have used different features in bio potential signals, this is because of the characteristics of the EEG signal. However, there are different frequency band of EEG signals, such as alpha beta, delta, and gamma. We have work out of our EEG dataset based on time domain and frequency domain features extraction directly from the signal. The important time domain features such as, maximum value, mean value, standard deviation, skewness, kurtosis etc., were extracted from raw EEG datasets. These are described below:

1) Mean Value:

The range of EEG potential is microvolts with time. The mean of EEG is constant and small values which change potentially. Definition of mean as described based on EEG signal is given by:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Where *x* is EEG signal and N is total number EEG data.

2) Standard Deviation:

Standard deviation is a measurement to find the amount of variation or dispersion from the average. To determine standard deviation of EEG signal, we have considered low and high variation. The standard deviation can be expressed as

$$std = \sqrt{\frac{\sum(x-\mu_x)^2}{N-1}}$$

3) Skewness:

It is a measurement of the asymmetry of the probability distribution of an EEG signal and it's mean. The mean value can be positive, negative or undefined. The mean values can be zero based on completely symmetrical distribution and some nonzero values are asymmetrically distribution in respect of the baseline.

$$Skew = \frac{\sum_{i=1}^{N}(x_i-\mu_x)^3}{(n-1)\sigma_x^3}$$

4) Kurtosis:

It's the measurement of the peakness of the probability distribution of EEG signal. The signal contains transient spikes, isolated high-voltage wave group. The kurtosis of it presents high positive values or negative values. These values are observed when EEG with high or low frequency and amplitude modulation is analyzed. The moment co-efficient can be written as

$$Kurt = \frac{\sum_{i=1}^{N}(x_i-\mu_x)^4}{(n-1)\sigma_x^4}$$

a) Coiflet Transform:

In the proposed method, the Coiflet wavelet was selected as the wavelet basis function. This wavelet exists under the name of Coiflets, but it is indeed constructed by I. Daubechies at the request of R. Coifman. Therefore, although Coiflet wavelet and well-known Daubechies wavelet are similar in a certain level, the Coiflet wavelet was indeed different in that it was constructed with vanishing moments not only for wavelet function Ã(t), but also for scaling function Á(t). In this way of design, the scaling function of the Coiflet will exhibit interpolating characteristics, which also implies that this wavelet allows a very good approximation of polynomial function at different resolution.

In the Coiflet wavelet formulation process, the following equations must be satisfied

$$\int dx\, x^l \psi(x) = 0 \qquad \text{for} \quad l = 0,\ldots,L-1$$

$$\int dx\, \phi(x) = 1$$

$$\int dx\, x^l \phi(x) = 0 \qquad \text{for} \quad l = 1,\ldots,L-1$$

where L is the order of Coiflet wavelet function, Á(x) is a scaling function associated to Ã(x). For a Coiflet of order 6, Figs. 1(a) and (b) plot the corresponding scaling function and wavelet basis function. These functions were shown to be smoother and more symmetric than Daubechies wavelet [. This observation indicates that at different resolution, the approximation of polynomial functions can be better achieved. Furthermore, the symmetry property of the Coiflet is desirable in the signal analysis work due to the linear phase of the transfer function. As for the comparison with Morlet wavelet that was also tested in our laboratory [16—18], although the Coiflet method is less flexible in visualizing any frequency of interest, its discrete form is useful for the digital implementation. These benefits consolidate the utilization of Coiflet wavelet transform for this study.

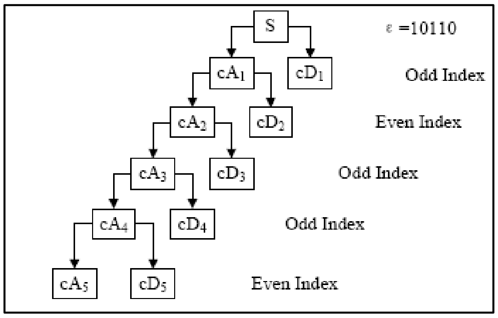
b) Stationary Wavelet Transform (SWT):

The DWT suffers from time variant property. This means that the DWT of a translated version of a signal X is not the translated version of the DWT of X.

D. ε-decimated DWT :

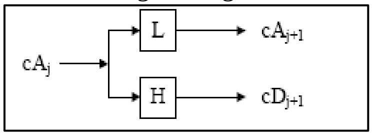
There exist a lot of slightly different ways to handle the DWT. The decimation in DWT retrains even indexed elements, which is where the time variant problem lies in. The decimation could be carried out by choosing odd indexed elements instead of even indexed elements. The choice of even or odd concerns every step of the decomposition process. If we perform all the different possible decompositions of the original signal for a given maximum level J, then we will have 2J different decompositions.

Suppose εj=1 or 0 denotes the choices of odd or even indexed elements at step j. Then, every decomposition is labeled by a sequence of 0s and 1s, namely, ε=ε1ε2…εJ. This transform is called the ε-decimated DWT. A graphical example of ε=10110 is shown in Fig. 4



1) 1D SWT:

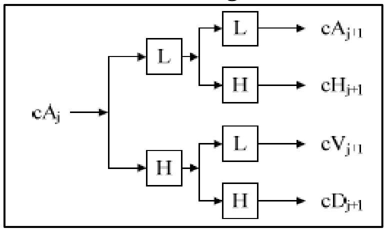
The SWT can calculate all the ε-decimated DWT for a given signal at one time. More precisely, for level 1, the SWT can be obtained by convolving the signal with the appropriate filters as in the DWT but without down sampling. Then the coefficients of the approximation and detail at level 1 are the same as the signal length.



The general step j convolves the approximation coefficients at level j-1, with appropriate filters but without down sampling, to produce the approximation and detail coefficients at level j. The schematic diagram is shown in Fig. 5.

2) 2D SWT:

The algorithm of 1D SWT can be easily extended to the 2D case. Fig. 6 shows the schematic diagram of 2D SWT.



3) Walsh Hadamard Transform (WHT):

Linear image transforms [25] such as: Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Walsh Hadamard Transform (WHT) and Karhunen-Loeve transform (KLT) have been used in various image/video processing applications due to their energy compaction, entropy, flexibility, robustness and performance. Through analyzing or classifying the image frequencies or coefficients of the linear image transforms, edges of the image can be detected. The kernels or basis of the image transforms used for extracting edges can also be used to compute a set of energy measures which could characterize the local texture properties of a given region in an image. Hence, the kernels can extract the local texture property.

Among the linear image transforms, WHT is very attractive due to the simplicity of its implementation and to its properties which are similar to other transforms. Also, the computational efficiency of WHT makes it very attractive image processing directly in the transform domain since the components of the basis vectors are orthogonal and have only binary values (−1 or +1). WHT is a suboptimal, non-sinusoidal, orthogonal transform and it is used in many different applications, such as filtering, processing speech and medical signals, etc. More specifically, this transformation is used for astronomical signal/image processing, coding and filtering operations. WHT is well known for its simple and fast transformation.

The WHT is represented as a matrix and constructed from the WHT Matrix (WHTM) [25]. An WHTM is defined as a set of N rows, denoted Wj , for j = 0, 1,…, N − 1, which have the following properties i) Wj takes on the values +1 and −1, ii) Wj[0] = 1 for all j, iii) Wj has exactly j zero crossings, for j = 0, 1,…, N−1. The size of a transform matrix is generally a power of 2. The matrix exists when N>2 and (N mod 4) = 0. The sequence ordered WHTM of order 4 is

$$For\ \ N = 4 : \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

Each row of WHTM is called a 1D Walsh Hadamard Basis vector. In general, basis vectors are orthogonal (dot product of any two of them is zero) and orthonormal (dot product of each with itself is 1). Using the tensor of 1-D basis vectors of WHTM, the 2D WHT kernels (WHTK) or basis images are generated by multiplying the corresponding row and columns as given below:

$$Tensor\ product\ of\ (2,2)\quad \begin{array}{c} \\ 1 \\ 1 \\ -1 \\ -1 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 1 & -1 & -1 \end{array} \\ \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \end{array}$$

The above matrix is the tensor product of the 2nd row and 2nd column basis vector of WHTM. Each tensor product is a Basis image or Kernel of the WHT. Given an N × N block of pixels (where N must be a power of 2, N = 2n), its WHT is obtained by projecting the block over the WHT Kernels. However, these kernels can be represented as vectors forming the Basis Vectors of WHT and are represented in vector space as V = v1, v2,...,v16 named from left to right of the WHT Kernel given in Fig. 1. The WHT kernels (or basis images) for all basis vector of WHTM of order 4 is shown in Fig. 1 and the kernels below the diagonal representation are transpose of the kernels above the diagonal. From here onwards, the term basis vector represents the basis vector of WHT.

E. KNN Ensemble Classification:

1) Ensemble Learning

Ensemble learning is an effective and increasingly adopted technique that combines multiple learning algorithms to improve overall performance. Ensemble learning consists of the following four main parts.

- Bagging: bagging, based on the majority-voting concept, utilizes randomly selected training data subsets, for training a dissimilar base learner of a similar manner.
- AdaBoost: adaboost is a famous member of the boosting approach. It creates base classifiers via sequential bootstrap samples, gained by weighting the training transactions via numerous iterations. Weighting is adjusted through mis classification related to the base classifier.
- Stacking: stacking combines different learning algorithms, in order to achieve higher prediction accuracy.

2) K-nearest neighbor

K-nearest neighbor (k-NN) is a simple classification model that exploits lazy learning. It is a supervised learning algorithm, which classifies new instance queries based on the majority of the k-nearest neighbor category. Calculating the minimum distance between the query instance and the training sample approximates the k-NN category. The k-NN prediction of the query instance is determined based on majority voting of the nearest neighbor category.

IV. EXPERIMENT AND RESULTS

To examine the effectiveness of the proposed method and to prove its advantage over the othert methods, the proposed method is experimented over the Bonn University dataset and it is evaluated by performance criteria. In this section, we focus mainly on dataset, performance measures and comparison of the proposed method with the existing method.

A. Dataset and Evaluation Criterion

To enable comparison with other other techniques, the proposed method was tested on the Matlab2018b software evaluation data. There are both epileptic and non-epileptic sequences. n addition to this, in order to compare the proposed method with recent related works, the proposed method is tested with some software listed in Table I as follows:

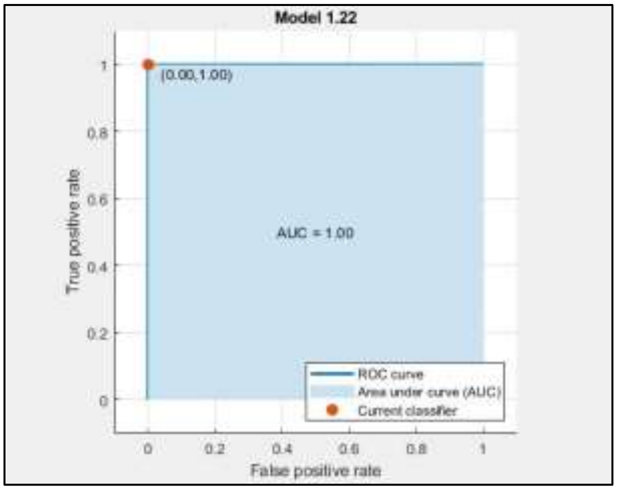


Fig. 1: Showing Roc Cure of the Process Method

Table - 1
Comparison Results of Existing Feature Extraction, Classifications and Performance Techniques for EEG Dataset.

Sr.No.	*Classifiers Name*	*Accuracy (%)*
1	Medium Gaussian SVM	99.8
2	Coarse Gaussian SVM	99.1
3	Fine KNN	99.4
4	Medium KNN	98.7
5	Coarse KNN	94.9
6	Cosine KNN	95.3
7	Cubic KNN	97.0
8	Weighted KNN	98.7
9	Fine tree	99.8
10	Medium tree	99.8
11	Coarse tree	99.8
12	Logistic regression	99.8
13	Linear SVM	**99.9**
14	Cubic SVM	99.4
15	Quadratic SVM	99.8
16	Fine Gaussian SVM	99.0
17	Ensemble boosted tree	50
18	Ensemble bagged tree	99.8
19	Ensemble subspace discriminant	90.9
20	Ensemble subspace KNN	**99.9**
21	Ensemble RUS boosted tree	50

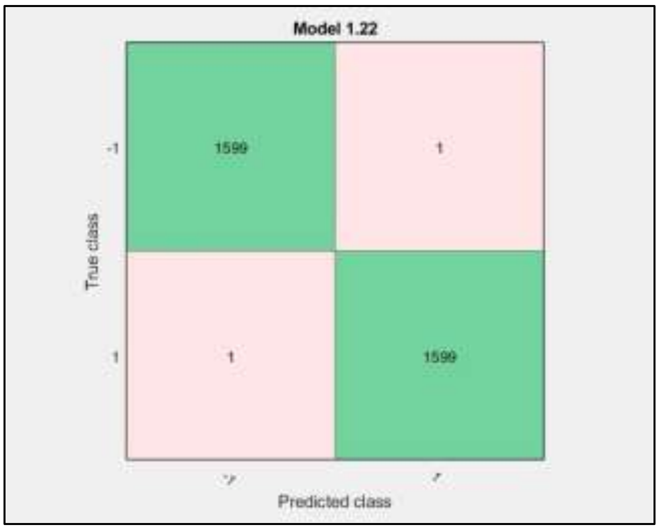


Fig. 2: showing 99% accuracy of process algorithm

V. CONCLUSION

This paper presented an EEG data classification algorithm to epileptic and non-epileptic by using existing dataset which, based on a large number of features extracted after wavelet transform .We have studied the existing literature and found that very few approaches can successfully classify the EEG signal with less number of samples. It is challenging problem to be solved using machine learning techniques in coming months an efficient algorithm for EEG signal classification can be developed. Therefore, the conclusion is that the proposed algorithm can be used to classify EEG signals with 99% accuracy.

REFERENCES

[1] Hosseini M-P, Nazem-Zadeh M-R, Pompili D, Jafari-Khouzani K, Elisevich K, Soltanian-Zadeh H." Comparative performance evaluation of automated segmentation methods of hippocampus from magnetic resonance images of temporal lobe epilepsy patients." Med Phys 2016;43:538–53.
[2] Hosseini M-P, Hajisami A, Pompili D. Real-time epileptic seizure detection from EEG signals via random subspace ensemble learning. In: 2016 IEEE international conference on autonomic computing (ICAC). IEEE; 2016. p. 209–18.
[3] Kuzum D, Takano H, Shim E, Reed JC, Juul H, Richardson AG, de Vries J, Bink H, Dichter MA, Lucas TH, Coulter DA, Cubukcu E, Litt B. Transparent and flexible low noise graphene electrodes for simultaneous electrophysiology and neuroimaging. Nat Commun 2014;5.
[4] Jayapandian C, Chen C-H, Dabir A, Lhatoo S, Zhang G-Q, Sahoo SS. Domain ontology as conceptual model for big data management: application in biomedical informatics. In: International conference on conceptual modeling. Springer; 2014. p. 144–57.
[5] Pagán J, Risco-Martín JL, Moya JM, Ayala JL. Modeling methodology for the accurate and prompt prediction of symptomatic events in chronic diseases. J Biomed Inform 2016;62:136–47.
[6] Ombao H, Lindquist M, Thompson W, Aston J. Handbook of neuroimaging data analysis; 2016.
[7] Cui L, Sahoo SS, Lhatoo SD, Garg G, Rai P, Bozorgi A, Zhang G-Q. Complex epilepsy phenotype extraction from narrative clinical discharge summaries. J Biomed Inform 2014;51:272–9.
[8] Combi C, Pozzani G, Pozzi G. Telemedicine for developing countries: a survey and some design issues. Appl Clin Inform 2016;7:1025–50.
[9] Zhou W, Liu Y, Yuan Q, Li X. Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG. IEEE Trans Biomed Eng 2013;60:3375–81.
[10] Khan YU, Gotman J. Wavelet based automatic seizure detection in intracerebral electroencephalogram. Clin Neurophysiol 2003;114:898–908.
[11] Ghosh-Dastidar S, Adeli H, Dadmehr N. Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE Trans Biomed Eng 2007;54:1545–51.
[12] Conradsen I, Beniczky S, Hoppe K, Wolf P, Sorensen HB. Automated algorithm for generalized tonic–clonic epileptic seizure onset detection based on SEMG zero-crossing rate. IEEE Trans Biomed Eng 2012;59:579–85.
[13] Hopfeng R, Kasper BS, Graf W, Gollwitzer S, Kreiselmeyer G, Stefan H, Hamer H. Automatic seizure detection in long-term scalp EEG using an adaptive thresholding technique: a validation study for clinical routine. Clin Neurophysiol 2014;125:1346–52.
[14] Pachori RB, Patidar S. Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions. Comput Methods Programs Biomed 2014;113:494–502.
[15] Hassan AR, Siuly S, Zhang Y. Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating. Comput Methods Programs Biomed 2016;137:247–59.
[16] Rana P, Lipor J, Lee H, Van Drongelen W, Kohrman MH, Van Veen B. Seizure detection using the phase-slope index and multichannel ECOG. IEEE Trans Biomed Eng 2012;59:1125–34.
[17] Orosco L, Correa AG, Leber EL. Epileptic seizures detection based on empirical mode decomposition of EEG signals. Manag Epilepsy Res Results Treat 2011:1.
[18] Gotman J. Automatic seizure detection: improvements and evaluation.Electroencephalogr Clin Neurophysiol 1990;76:317–24.
[19] Gabor AJ. Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies. Electroencephalogr Clin Neurophysiol 1998;107:27–32.
[20] Wang Y, Qi Y, Zhu J, Zhang J, Wang Y, Pan G, Zheng X, Wu Z. A cauchy-based state-space model for seizure detection in EEG monitoring systems. IEEE Intell Syst 2015;30:6–12.
[21] Becq G, Kahane P, Minotti L, Bonnet S, Guillemaud R. Classification of epileptic motor manifestations and detection of tonic–clonic seizures with acceleration norm entropy. IEEE Trans Biomed Eng 2013;60:2080–8.
[22] Do Valle BG, Cash SS, Sodini CG, Low-power. 8-channel EEG recorder and seizure detector ASIC for a subdermal implantable system. IEEE Trans Biomed Circ Syst 2016;10:1058–67.
[23] Salam MT, Sawan M, Nguyen DK.A novel low-power-implantable epileptic seizure-onset detector. IEEE Trans Biomed Circ Syst 2011;5:568–78. Altaf MAB, Yoo J. A 1.83 J/classification, 8-channel, patient-specific epileptic seizure classification soc using a non-linear support vector machine. IEEE Trans Biomed Circ Syst 2016;10:49–60.
[24] Mirzaei M, Salam MT, Nguyen DK, Sawan M. A fully-asynchronous low-power implantable seizure detector for self-triggering treatment. IEEE Trans Biomed Circ Syst 2013;7:563–72.
[25] Yadav R, Shah A, Loeb JA, Swamy M, Agarwal R. Morphology-based automatic seizure detector for intracerebral EEG recordings. IEEE Trans Biomed Eng2012;59:1871–81.