Biomedical Signal Processing and Control 41 (2018) 233–241

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

A novel seizure diagnostic model based on kernel density estimation and least squares support vector machine

Mingyang Li, Wanzhong Chen*, Tao Zhang

College of Communication Engineering, Jilin University, Changchun, 130012, China

ARTICLE INFO

Article history:
Received 19 May 2017
Received in revised form 2 November 2017
Accepted 10 December 2017

Keyword:
EEG
WPT
Kernel density estimation (KDE)
LS-SVM

ABSTRACT

The automated system can be an effective tool for assisting neurologists in seizure detection. However, most of the existing methods are failed to trade off the effectivity and computation cost, which is not appropriate for on-line application. In this research, we propose a novel method for dealing with 3-class electroencephalogram (EEG) problem, based upon kernel density estimation (KDE) and least squares support vector machine (LS-SVM). The filtered EEG is decomposed into several sub-bands by wavelet packet transform (WPT), then KDE is explored to calculate the corresponding probability density. Five parameters are employed for EEG representation: the maximum (Max), the skewness (Ske), the kurtosis (Kur), the energy (En), and the central moment (CM). And significant features selected by Analysis of Variance (ANOVA) are fed to LS-SVM for pattern recognition. Furthermore, eight types of wavelet bases and four well-known functions are considered for feature extraction. Experimental results show that our approach has achieved satisfactory and comparable results for all validation methods when configured with coiflet of order 1 and uniform kernel. The highest accuracy of 10-fold cross-validation and standard 50-50 methodology is 99.40% and 99.60% with 27 and 26 features, respectively. As compared to previous literature, our proposed scheme is more suitable for diagnosis of epilepsy with higher accuracy and less number of feature that can be extracted with less computational cost. Overall, the advantages of high accuracy, easy implementation and low computational consumption have made this technique a suitable candidate for extensive clinical deployment.

1. Introduction

Epilepsy is a chronic neurological disorder characterized by the sudden and excessive neural discharges in the brain [1,2]. It has been estimated that approximately one in every 100 people worldwide are suffering from epilepsy [3,4]. Until recently, epilepsy has become a global problem in the public healthy, which severely affected the patients' life quality, study, and working abilities. Electroencephalography (EEG) is an important clinical tool that is used for epileptic detection as it is a condition related to the brain activity [5,6]. However, visual inspection of the EEG recordings by experienced neurologist is a very time consuming, cumbersome and subjective task [7]. Hence, the automatic seizure detection technology holds great significance and prospect in clinical implications, which can help doctors confirm their initial diagnosis as well as develop suitable treatment of patients with the personalized.

Various methods addressing feature extraction have been proposed for the purpose of epileptic seizure detection. These methods can be broadly summarized as four categories namely time domain analysis [8], frequency domain analysis [9], time-frequency domain analysis [10] and non-linear dynamics analysis [11]. There has been an increasing interest in the study of seizure detection using wavelet transform (WT) and wavelet packet transform (WPT) in recent years. Subasi [10] used statistical features over the set of the discrete wavelet transform (DWT) coefficients for the classification of epileptic EEG signals. Ocak [12] developed a system for two-class epilepsy detection based on DWT and approximate entropy (ApEn). Kumar et al. [13] have employed the DWT based fuzzy approximate entropy as a feature for automated seizure detection using support vector machine (SVM). Martis et al. [14] has proposed an epileptic EEG classification method using WPT and non-linear parameters. Acharya et al. [15] modified the feature extraction with the use of WPT along with principal components analysis (PCA). Although both DWT and WPT are popular in feature extraction of EEG signals, WPT is able to offer better partial characteristics and analyze the information both in high and low frequencies. So in the paper, we adopt WPT for signal processing.

* Corresponding author at: College of Communication Engineering, Jilin University, Ren Min Street, 5988, Changchun, China.
E-mail address: chenwz@jlu.edu.cn (W. Chen).

https://doi.org/10.1016/j.bspc.2017.12.005

As noted in earlier works that most methods are likely to extract non-linear or statistical features in the DWT or WPT domain. However, the non-linear algorithms will incur increased computational time requirement complexity. And statistical features extracted in only time-frequency domain are limited in their subtle characteristics representation. The constraints have made them hard to be implemented in real time application. Hence, a hybrid method based on the kernel density estimation (KDE) and WPT are proposed in this research. As far we are aware of, there is no study in the literature related to the KDE in the diagnosis of epilepsy. The KDE is a non-parametric way for estimating probability density function (PDF) of signals, which is extensively used in image processing because of its discriminatory power and computational simplicity. In this regard, the KDE is deployed in combination with WPT and five statistical features are derived in the KDE-WPT domain. By this means, both the statistical properties and transient changes can be captured and localized. Particularly, we have not only exploited the use of KDE in seizure detection but also considered the influence brought by different wavelet bases and kernel functions through experimental evaluation.

This study proposes a new seizure detection method that is different from the approaches presented in the previous studies. Fig. 1 has shown a block diagram of the proposed technique. The data used in this work is divided in two parts, one is for model building and the other is for model testing. As seen in Fig. 1, the filtered EEG are subjected to WPD for 4-level decomposition. Then both the raw EEG and the WPD coefficients are analyzed by the KDE for corresponding PDF, from which five statistical features such as the maximum (Max), the skewness (Ske), the kurtosis (Kur), the energy (En), and the central moment (CM) are obtained. In order to eliminate redundant information, highly significant features were selected using the Analysis of Variance (ANOVA) test. And least squares support vector machine (LS-SVM) is employed as the classifier for deciding the class of the input features. The rest of the paper is designed as follows: Section 2 provides the description of the data and the proposed methodology. And Section 3 presents the experimental results. A discussion of our approach is given in Section 4, followed by the summary of this research in the last Section.

2. Materials and methods

2.1. EEG dataset

The data used in this paper is obtained from the Department of Epileptology, University of Bonn. More detailed information about the data is provided in Ref. [16] including the acquisition process, which are not mentioned to keep this paper reasonably concise. The whole database consists of five sub-sets denoted as Z, O, N, F, and S, each containing 100 single-channel EEG segments of 23.6 s duration. These EEG signals are recorded at a sampling rate of 173.61 Hz using a 128-channel amplifier system with an average common reference. In this study, sub-sets Z, F, S are used for further analysis since the case of Z, F and S (simplified as Z-F-S) has become a classical and complicated classification problem and attracted the great attention of researchers in recent years. Set Z consist of surface EEG segments collected from five healthy volunteers with their eyes open. The segments in both F and S were obtained from sick volunteers with electrodes placed in the epileptogenic zone. Group F was recorded during seizure-free interictal trials, while group S was measured during seizure activity. The EEG are classified into three different classes namely normal intervals (Z), inter-ictal intervals (F) and ictal intervals (S). Sample EEG signals from the three sets are plotted in Fig. 2.

2.2. KDE based feature extraction

2.2.1. WPT decomposition

WPT, introduced by Coifman and Wickerhauseran [17], is known as the further development of discrete wavelet transform (DWT). WPT has attracted increasing attention because of its ability in providing more flexible time-frequency decomposition, especially in the higher frguency region [18]. Since DWT decomposes only approximate information into each successive level, the analysis of WPT is more detailed, and has more accurate partial analysis ability.

In the case of WPT, both detail and approximation coefficients are generated for further decomposition at each level. By decomposing the original signal, one approximation coefficient and one detail coefficient are obtained at the first level. Similarly, these two components are further decomposed into four coefficients at level 2, and so on. Due to the binary tree structure, WPT offers better resolution in comparison with DWT. Hence, WPT is a superior approach to preciously reveal and localize transient features in epileptic data. The choice of decomposition levels and suitable wavelet function directly affects WPT processing results. The number of decomposition levels is thus fixed to 4 in this present paper, the same number adopted by Zhang et al. [19]. As well, we have performed an empirical study with different types of wavelets aiming at finding the appropriate wavelet bases for particular applications.

2.2.2. KDE analysis

Kernel density estimation (KDE) is a well known technique that is widely used in the field of statistics and pattern recognition [20]. As a non-parametric model, KDE provides a smooth, continuous, and differentiable density estimate without assuming any specific underlying distribution. Owing to these advantages, we have deployed KDE into EEG analysis, which can not only describe a time series but also capture the features of distribution density. Let x_i, ($i=1,\ldots,n$) be an independent sequence drawn from an arbitrary probability distribution. And the form of KDE can be given:

$$\hat{f}(x)=\frac{1}{Nh}\sum_{i=1}^{N}G(\frac{x-x_i}{h}) \tag{1}$$

Where f denotes the probability density function (PDF), N is the sample number, G is the kernel function and h is the bandwidth. The kernel function is an important factor for final joint distribution and the four popular functions are summarized:

1) *Uniform kernel* (*Un*)

$$G(x)=\begin{cases}\frac{1}{2}, |x|\le 1\\ 0, \text{other}\end{cases} \tag{2}$$

2) *Triangle kernel* (*Tr*)

$$G(x)=\begin{cases}1-|x|, |x|\le 1\\ 0, \text{other}\end{cases} \tag{3}$$

3) *Epanechnikov kernel* (*Ep*)

$$G(x)=\begin{cases}\frac{3}{4}(1-x^2), |x|\le 1\\ 0, \text{other}\end{cases} \tag{4}$$