IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 6, 2019, accepted November 26, 2019, date of publication November 29, 2019,
date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2956768

Automated Change-Point Detection of EEG Signals Based on Structural Time-Series Analysis

GUANGYUAN CHEN, GUOLIANG LU, (Member, IEEE), WEI SHANG, AND ZHAOHONG XIE

Key Laboratory of High-efficiency and Clean Mechanical Manufacture of MOE, National Demonstration Center for Experimental Mechanical Engineering Education, School of Mechanical Engineering, Shandong University, Jinan 250061, China

Corresponding authors: Guoliang Lu (luguoliang@sdu.edu.cn) and Wei Shang (wshang85@aliyun.com)

This work was supported in part by the Shandong Provincial Natural Science Foundation, China, under Grant ZR2019MEE063, and in part by the Fundamental Research Funds of Shandong University under Grant 2018JC010.

ABSTRACT Automated change-point detection of EEG signals is becoming essential for the monitoring of health behaviors and health status in a wide range of clinical applications. This paper presents a structural time-series analysis to capture and characterize the dynamic behavior of EEG signals, and develops a method to detect the EEG change points. For a given EEG signal, the proposed method is operated as follows: 1) a sub-band pass filter is fist designed to capture those frequency components that can characterize the dynamic behavior of the data, and the so-called power spectrum is extracted as the EEG features; 2) together with a sliding-window technique, an automatic 'segment-to-segment' analysis of EEG signal, is developed with a null hypothesis testing for decision making. In particular, the main challenge of the proposed method is to design an appropriate distance metric that is compatible with our considered data/problem. To achieve this end, we first collect a variety of metrics from other areas that would be potentially available for our problem, and then compare them for the considered EEG change point detection. Experiments are conducted on two different data sets. Results show the Bhattacharyya distance achieves the best detection result among all investigated metrics. Meanwhile, comparison with state-of-the-arts demonstrates the effectiveness of the method in real applications.

INDEX TERMS Change detection, distance metric, EEG monitoring, similarity metric, time series analysis.

I. INTRODUCTION

Electroencephalogram (EEG) is a powerful and effective tool to record and comprehend complex activities of the brain [1]. Recent advances in sensing, communication, computation, visualization, etc., provide the ability to acquire large streams of EEG signals, offering an unprecedented opportunity to detect change-points of EEG signals at an early stage [2]–[4] and substantially improve health assessment as well as many other neurological diseases related health-care and surgery problems [5]–[7].

With such a motivation, pattern recognition techniques and rule-based inferences have been proposed with an assumption of the availability of massive histories of 'irregular' EEG recordings [8], [9]. In real-life scenarios, however the irregular patterns do not occur frequently and furthermore are individual-specific, which leads to the difficulty of collection of sufficient histories for such data [10].

Some previous researches have treated the problem as a novelty detection task [11], [12]. Novelty detection recognizes new EEG inputs that differ in some way from those that are usual under normal status. This paradigm overcomes one important limitation of competing pattern recognition and rule-based methods, i.e., the need for a pre-collection and labeling of irregular EEG data histories [13]. However, there are still some important issues that should be considered in such novelty detection methods. For example, some methods are quite heuristic and require manual selection of parameters, and meanwhile, the optimisation of a pre-defined number of parameters that define the structure of the model should be optimized [14], [15]. Therefore, further development and commercialization of this method is presently inhibited in

The associate editor coordinating the review of this manuscript and approving it for publication was Jason Gu.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

their real implementation. In this regard, continuous efforts have been devoted with various enhancements of this kind of methods until the present day [12], [16].

Along the line of ongoing research on novelty detection of EEG signals, this paper presents a new method based on structural time-series analysis. The flowchart of the method is shown in Fig. 1. Its execution procedure is given as follows. For a given EEG signal, a sub-band pass filter is first designed based on domain knowledge, in order to capture the frequency components that can characterize the dynamic behavior of the EEG data. Power spectrum is then extracted as the feature for EEG data representation. Using the sliding-window technique, the similarity of current segment with respect to the past EEG histories is measured/quantified. Change decision is finally made by testing a null hypothesis. The method is operated continuously and periodically, such that the collected EEG signal can be monitored and checked with an online and realtime manner. With respect to existing methods [9], [11], [17], the proposed approach presents the following advantages: 1) it is computationally efficient enough to be used for online monitoring since no data modeling as well as parameter estimation is needed; 2) it is self-conducted and unsupervised since any prior training of a classifier or detector is not required.

The proposed method is validated on two data sets: one with data from the public Bern-Barcelona EEG database and another with data collected using our system setup. Comprehensive experimental results indicate good potentials of the proposed method in real applications.

To summarize, our main contributions in this paper are described as follows.

1) A new scheme for early change detection in EEG signals based on structural time series analysis (Section III). It is the first and essential step to characterize the dynamical behavior of time series in order to produce reliable and accurate change detection result. Current methods using global statistics such as mean, root mean square (RMS), Kurtosis, etc., are not well positioned to effectively abstract information from EEG signals with highly noise and artifact [18], [19]. And those methods based on support vector machine (SVM) and genetic algorithm (GA) require large amount of data and consume large time to train [20], [21]. In comparison, our method relies on a specific metric time-sequentially comparing two consecutive data segments and periodically analyzing their similarities for making final result. Thus, the proposed method has obvious advantages such as more robustness to noise and a lower computational complexity, that is $O(n \log n)$ where n denotes the number of data points in a given EEG signal.
2) Evaluation of a variety of metrics for EEG similarity quantification (Section III-B). Designing an appropriate similarity metric, that is compatible with the considered data, is an important aspect to design such change detection systems. However, it is impossible to apply existing metrics directly to the EEG signals due to the domain specificity. That is to say, the metric used for EEG signal processing still needs to be clarified. To achieve this end, a variety of most popular and state-of-the-art metrics are taken from other areas, and modified/extended if necessary to incorporate them with the EEG change detection. Impacts of different metrics on change detection results are evaluated and investigated, where the Bhattacharyya Distance (BD) is demonstrated outperforming performances than other competitors.

The rest of this paper is organized as follows. Section II describes the feature extraction. Section III formulates the EEG change detection problem and investigates several potentially-available similarity metrics. The algorithm of proposed framework is shown in Section IV, followed by experimental validation in Section V. Section VI finally concludes this paper with some remarking points and shows the future work.

II. FEATURE EXTRACTION

EEG signals are normaly considered as complex and non-stationary, and meanwhile data redundant. Thus it has been a common practice in EEG signal analysis to extract accurate and reliable features from raw EEG data. The most popularly used methodology is the time-frequency features as discussed extensively in many previous works, e.g., [22]–[24]. A well-established mechanism of time-frequency analysis is to first confirm a specific sub-band based on domain knowledge, and subsequently calculates its statistics to extract features. Effective and accurate feature extraction provides a precise understanding of the dynamic behavior of EEG time series, supporting a successful detection of possible changes. If there is a change happening, the extracted features will change in comparison to normal EEG status in a determined way.

In this section, we introduce the employed feature extraction method, which includes: 1) pre-determination of sub-bands, and 2) power spectrum calculation. Here, it is noted that, our proposed method can also be applied to other features as long as the extracted feature can reflect the dynamic characteristics of monitored EEG data. However, since the focus of this paper is on the detection mechanism, we do not provide additional discussion on feature extraction. The interested reader can refer to [25], [26] for this issue.

A. PRE-DETERMINATION OF SUB-BANDS

The happening of neurogenic disorder is a phenomena that can be interpreted as a change from normal EEG status to another. The main frequency components of indication of a neurogenic disorder happening are: δ, θ, α, β, γ [27]. That is, if a neurogenic disorder happens, the amplitudes of these frequencies change accordingly. Actually, although the majority of existing works (e.g., [28], [29]) are focusing on the frequency components from 0.1-30Hz that covers δ, θ, α, β, there is also an interest on frequency above 30Hz in EEG signal analysis [30]. Since the aim of this study is to explore a unify framework to detect the change point of EEG

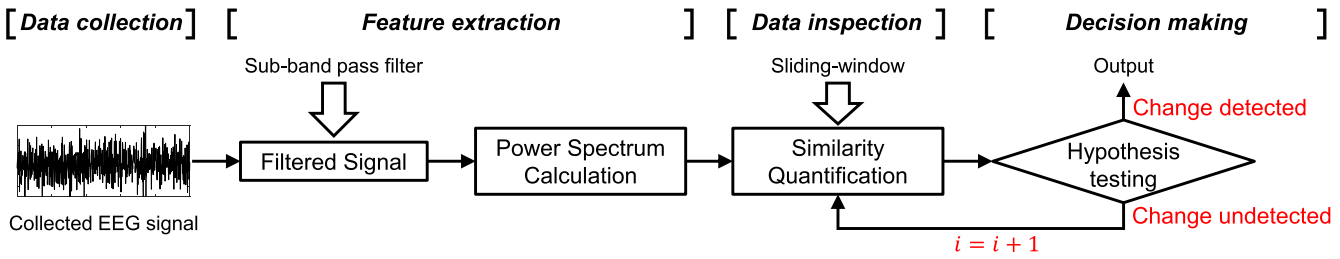


FIGURE 1. The flowchart of the proposed method.

signals, we do not make additional discussion on frequency selection, and alternatively set the monitoring frequency by a sub-band pass filter of [0.1, 70]Hz in the following of this paper. Here, it is also noted that, the effectiveness of the proposed method does not rely on the setting of monitoring frequency, interested readers can use the proposed method with respect to monitoring frequencies in their studies.

Fig. 2 illustrates the filtering process. The raw EEG data is filtered using the pre-determined sub-band pass filter. A notch filter is also used to remove the power line interference which is around 50Hz. As such, the resulting data can only contain frequency components from 0.1 Hz to 70 Hz.

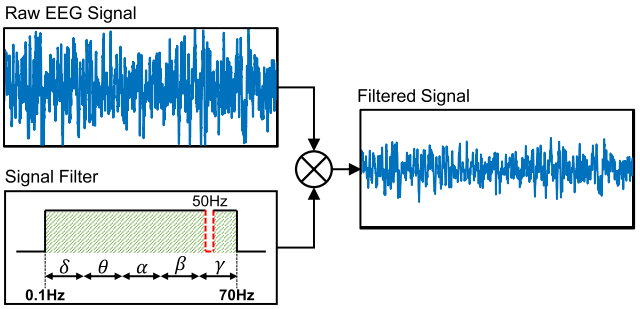


FIGURE 2. The sub-band filtering for raw EEG data.

B. POWER SPECTRUM CALCULATION

A commonly used methodology to analyse predetermined sub-bands is performing power spectrum analysis based on short-time Fourier transform (STFT) [31], [32]. Let us assume that the observed value of filtered EEG signal at time t has been denoted as $d(t)$. We may review the STFT calculation. The integral formulation is given as,

$$F(f,\tau) = \int_0^t d(t)\omega(t-\tau)e^{-j2\pi f\tau}dt \tag{1}$$

where $F(f,\tau)$ is the output of the transform at frequency f and at time τ. The proper window ω, also called the weighting function, can be selected according to the characteristic of processed signals with considering the amplitude resolution, frequency resolution, time resolution, etc. This size of the window is a compromise between two requirements: (a) it should not be sensitive to noise and (b) it should be able to accurately concentrate the detect in the window where the change is significant. If the size of the window is too narrow, the method will be very sensitive to noise. On the other hand, too wide window would not allow us to accurately localise the point where the change is evident. The size of the window used in this work is 0.2 seconds as a compromise of these two goals.

Second, let us consider the discrete situation where the transform is discrete in both time and frequency domains. When both the sampling frequency f_s and window length T are fixed, those related parameters are calculated accordingly as given in Table 1. The spectrum can be computed at discrete frequencies, i.e., $f_0 = k * \Delta f = k * 1/T$, where k is an integer and $0 \le k < Tf_s/2$. Thus, the discrete form of $F(f,\tau)$ can be given as,

$$F(k,m) = \sum_{n=0}^{N-1} d[n]w[n-m]e^{-jk\Delta fn} \tag{2}$$

where N is the number of frequency components, m indicates the time index. Subsequently, the power spectrum $\hat{P}(k,m)$ can be estimated using the periodogram method by,

$$\hat{P}(k,m) = \frac{1}{N}|F(k,m)|^2 = \frac{1}{N}|\sum_{n=0}^{N-1} d[n]w[n-m]e^{-jk\Delta fn}|^2 \tag{3}$$

An example is shown in Fig. 3. One can see that, the testing EEG signal includes a change that indicates a transit from normal to abnormal (see Fig. 3(a)). However, as seen in Fig. 3(c), the signal may has similar power spectrums in normal and abnormal status. This implies, although the power spectrum can be used as feature for EEG data representation, further analysis is still needed to detect the change-point. In the following, an analysis of EEG signals by means of statistical time-series analysis will be presented.

III. METHODOLOGY

This section first provides a common scheme for continuous monitoring of EEG signal based on the sliding-window technique, and then formulates the problem that needs to be considered.

A. PROBLEM FORMULATION

Using the feature extraction given above, the given monitoring EEG data stream D (up to the inspection time t) has been represented by a sequence of spectrums based on a sliding-window, i.e., $P = \{p_1, p_2, ..., p_t\}$ where p_i indicates the power spectrum extracted at time i.

TABLE 1. Parameters for STFT calculation.

Sampling interval	Sampling number	Frequency resolution	Frequency range
Δt	η	Δf	F
$1/f_s$	Tf_s	$1/T$	$0 - f_s/2$ Hz

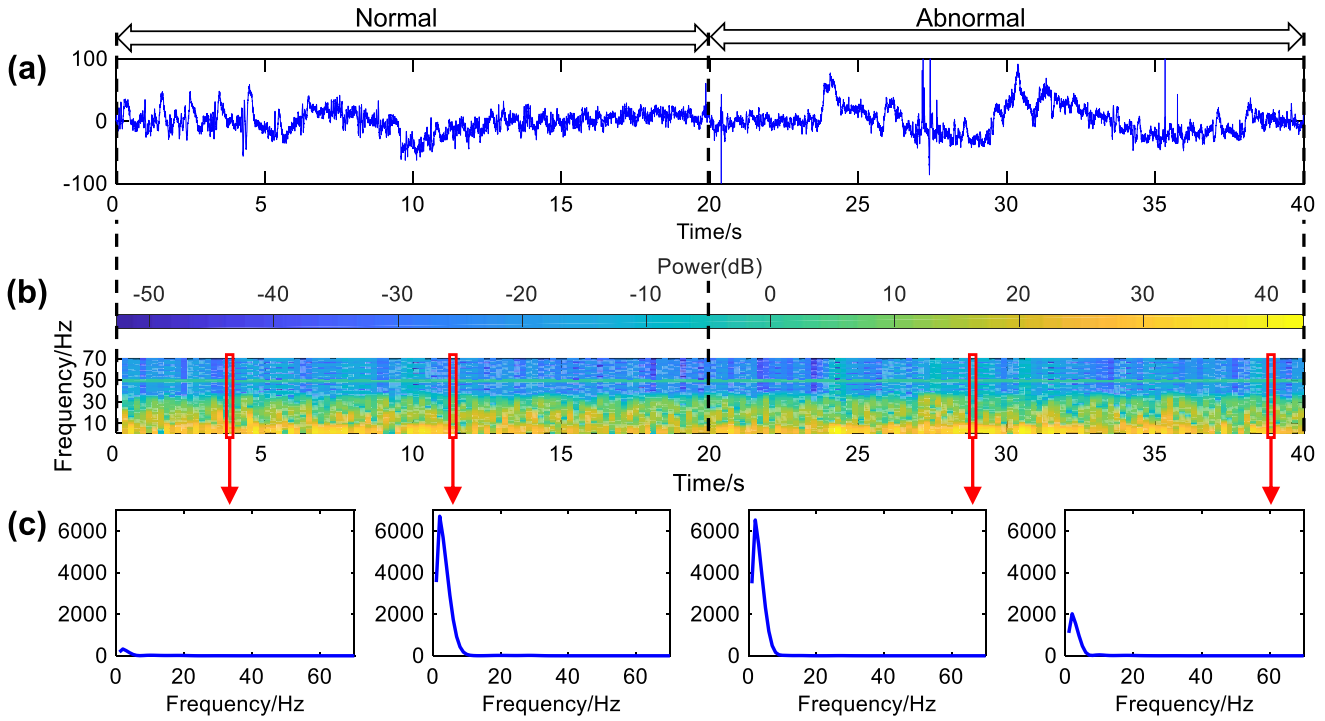


FIGURE 3. An example of STFT calculation where (a) a testing EEG signal including a change from normal to abnormal, (b) the computed power spectrum, and (c) power spectrums of several EEG data segments.

The problem of EEG change detection can be formulated as finding the change time that implies a transit from one data distribution D_1 to another D_2. This task is accomplished by a sliding window strategy including three steps given as below.

1) For an inspection time i, two candidate distributions are formed using the data within the sliding window, i.e., $D_1 : \{d_{i-L}, d_{i-L+1}, \ldots, d_i\}$ and $D_2 : \{d_i, d_{i+1}, \ldots, d_{i+L}\}$;
2) The STFT is used to calculated the power spectrum of the candidate data distribution with the predetermined sub-bands, i.e., $D_1 \rightarrow P_1 : \{p_1^1, p_2^1, \ldots, p_n^1\}$ and $D_2 \rightarrow P_2 : \{p_1^2, p_2^2, \ldots, p_n^2\}$
3) We can check the similarity of these two distributions with a specific score for the inspection time, which is denoted by $\mathcal{L}(P_1, P_2|i)$.

Time-sequentially performing the similarity inspection with an overlapping sliding window, we can obtain a time series similarity scores up to the time i, i.e., $\{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_i\}$. Then, the commonly used 3σ control chart with an assumption of Gaussian distribution is performed to make final result by testing a null hypothesis as,

$$\begin{aligned} H_0 &: \mid \mathcal{L}_i - \overline{\mathcal{L}}_{i-1} \mid < 3\sigma'; \\ H_1 &: \mid \mathcal{L}_i - \overline{\mathcal{L}}_{i-1} \mid \geq 3\sigma'; \end{aligned} \tag{4}$$

where H_0 means that no change occurs on the time i as long as $\mid \mathcal{L}_i - \overline{\mathcal{L}}_{i-1} \mid < 3\sigma'$, and H_1 indicates that a change occurs when $\mid \mathcal{L}_i - \overline{\mathcal{L}}_{i-1} \mid \geq 3\sigma'$. Here, $\overline{\mathcal{L}}_{i-1}$ and σ' are the mean and standard deviation of the Gaussian distribution respectively, and calculated by,

$$\overline{\mathcal{L}}_{i-1} = \frac{1}{i-1} \sum_{j=1}^{i-1} \mathcal{L}_j, \tag{5}$$

$$\sigma' = \sqrt{\frac{1}{i-1} \sum_{j=1}^{i-1} (\mathcal{L}_j - \overline{\mathcal{L}}_{i-1})^2}. \tag{6}$$

Here, it is worth mentioning that, there exist other alternatives such as Gaussian Mixed Model (GMM) [33] and other non-Gaussian assumptions [34] for change detection. However, since the focus of this paper is on the modeling, we do not investigate these alternatives in this study.

The above-given idea relies on an appropriate similarity score i.e. $\mathcal{L}$, to quantify the distance between two candidate data distributions. More specifically, when the difference between two data distributions is large, the value of employed distance will be large which implies a high probability that a change occurs, but when the difference between two data distributions is small, the value of measurement distance will

be small which implies a low probability that a change occurs. Considering the diversity of existing metrics, a specific similarity metric can only be feasible for certain domains due to the domain specificity as extensively discussed in the literature, e.g., [35]. In other words, it would be impossible to apply them directly to different types of data in other domains.

Hence, in this paper, we consider a variety of typical similarity metrics taken from other areas that would be potentially available for our usage, and modify/extend them if necessary to incorporate to our problem. The details will be given in the following section.

B. SIMILARITY METRIC

The similarity metric $\mathcal{L}$ is essential to report an accurate and reliable detection result and its construction normally relies on a specific distance metric, i.e., a greater value of distance indicates a smaller level of similarity. More importantly, for two given feature sequences of EEG recordings $\hat{P}_1$ and $\hat{P}_2$, the employed distance metric needs to satisfy several fundamental properties:

- Non-negativity, i.e., $\mathcal{L}\{\hat{P}_1, \hat{P}_2\} \geqslant 0$;
- Identity, i.e., $\mathcal{L}\{\hat{P}_1, \hat{P}_2\} = 0$ if and only if $\hat{P}_1 = \hat{P}_2$;
- Symmetry, i.e., $\mathcal{L}\{\hat{P}_1, \hat{P}_2\} = \mathcal{L}\{\hat{P}_2, \hat{P}_1\}$;
- Triangle inquality, i.e., $\mathcal{L}\{\hat{P}_1, \hat{P}_2\} \leqslant \mathcal{L}\{\hat{P}_1, \hat{P}_3\} + \mathcal{L}\{\hat{P}_3, \hat{P}_2\}$, where $\hat{P}_3$ is a third EEG recording that is not equivalent to both $\hat{P}_1$ and $\hat{P}_2$.

Here, one can note that, the distance metric for similarity quantification is not necessary to meet all of these properties especially the triangle inequality, under which such kinds of distance are called as non-metric distances [36].

Based on the above definition of distance metric, the similarity metric can also be confirmed as $\mathcal{L} \in [0, 1]$ with value of 1 if two compared EEG recordings are identical and 0 if nonidentical at all. In the following, we identified some typical similarity metrics with potentials to solving our problem by careful reviewing the relevant literature. In particular, during the identification, two following issues were considered.

- The metric should satisfy three properties of scalability, sensitivity and coverage;
- Among various metrics, we only pay attention to the ones which only calculate the similarity between two sequences with equal lengths.

Given two time series $P_1 = \{p_1(k)\}, k = 1, 2, \ldots, K$, and $P_2 = \{p_2(k)\}, k = 1, 2, \ldots, K$, we introduce and consider the following distance metrics to quantify their similarity.

1) EUCLIDEAN DISTANCE (ED)

Euclidean Distance is the most common metric that refers to the real distance between two points in space [37]. It can be calculated by

$$d^{(ED)} = \sqrt{\sum_{k=1}^{n}(p_1(k) - p_2(k))^2}. \tag{7}$$

Taking into account the characteristics of similarity metric described above, we use the reciprocal of $d^{(ED)}$ to represent the similarity as

$$\mathcal{L}^{(ED)} = \frac{1}{d^{(ED)}}. \tag{8}$$

2) PEARSON CORRELATION COEFFICIENT DISTANCE (PCCD)

Pearson Correlation Coefficient Distance was proposed by Pearson. It is a statistic used to reflect the degree of linear correlation between two series, with values between -1 and 1. The larger the value of it, the stronger the correlation of the two series [38]. It can be calculated by

$$d^{(PCCD)} = \frac{\sum_{k=1}^{K}(p_1(k) - \overline{p_1})(p_2(k) - \overline{p_2})}{\sqrt{\sum_{k=1}^{K}(p_1(k) - \overline{p_1})^2}\sqrt{\sum_{k=1}^{K}(p_2(k) - \overline{p_2})^2}}. \tag{9}$$

So the similarity defined by *PCCD* is then calculated by

$$\mathcal{L}^{(PCCD)} = |d^{(PCCD)}|. \tag{10}$$

3) SYMMETRIC KULLBACK-LEIBLER DIVERGENCE (SKLD)

The Kullback-Leibler Divergence can be used to measure the difference between two probability distributions, widely used in information retrieval and data science [39]. It can be calculated by,

$$D(P_1 \parallel P_2) = \sum_{k=1}^{K}(p_1(k)\log(\frac{p_1(k)}{p_2(k)})), \tag{11}$$

$$D(P_2 \parallel P_1) = \sum_{k=1}^{K}(p_2(k))\log(\frac{p_2(k)}{p_1(k)})), \tag{12}$$

but it is not strictly a distance measure, because of the asymmetry, which means that if P_1 is not equal to P_2, then $D(P_1 \parallel P_2)$ is not equal to $D(P_2 \parallel P_1)$. In order to solve the problem, symmetric Kullback-Leibler divergence is very popular in various statistical distance metrics [40] and is calculated by

$$d^{(SKLD)} = \frac{D(P_1 \parallel P_2) + D(P_2 \parallel P_1))}{2}. \tag{13}$$

Then the similarity can be obtained as

$$\mathcal{L}^{(SKLD)} = \frac{1}{d^{(SKLD)}}. \tag{14}$$

4) HELLINGER DISTANCE (HD)

The Hellinger Distance was first proposed by Hellinger in [41]. Hellinger Distance is used in probability and statistics to measure the similarity between two probability distributions, which belongs to f-divergence [41]. It can be calculated by

$$d^{(HD)} = \frac{1}{\sqrt{2}}\|\sqrt{P_1} - \sqrt{P_2}\|_2 \tag{15}$$

Thus, the similarity based on *HD* can be calculated as

$$\mathcal{L}^{(HD)} = \frac{1}{d^{(HD)}}. \tag{16}$$

5) KOLMOGOROV DISTANCE (KD)

The Kolmogorov Distance was introduced by Kolmogorov [42]. The statistical distance plays an important role in probability theory and hypothesis testing [43], and it is widely used to measure the difference between two probability distributions [44]. It can be calculated by

$$d^{(KD)} = \|P_1 - P_2\|_\infty \tag{17}$$

Thus, the similarity based on *KD* can be calculated as

$$\mathcal{L}^{(KD)} = \frac{1}{d^{(KD)}}. \tag{18}$$

6) BHATTACHARYYA DISTANCE (BD)

In the statistics, the Bhattacharyya Distance [45], that is similar to but different from the Hellinger Distance, measures the similarity of two discrete or continuous probability distributions. It is closely related to the Bhattacharyya coefficient, which measures the overlap between two statistical samples or populations [46]. The Bhattacharyya coefficient can be used to determine the separability of the classification used in the measurement of two samples that are considered relatively close, and is defined as

$$\mathcal{L}^{(BD)} = -\ln(BC(P_1, P_2)), \tag{19}$$

where $BC(X, Y)$ is the Bhattacharyya coefficient

$$BC(P_1, P_2) = \sum_{k=1}^{K} (\sqrt{p_1(k)p_2(k)}), \tag{20}$$

IV. ALGORITHM OF PROPOSED FRAMEWORK FOR EARLY CHANGE DETECTION IN EEG SIGNAL

Based on the methodologies described in previous sections, the algorithm of proposed EEG early change detection system includes the following five steps:

1) Collect EEG signal in a continuous manner;
2) Divide the streaming data into individual segments with a non-overlapping window;
3) Perform feature extraction for each EEG signal segment;
4) Measure the similarity between two neighboring EEG segments, and test a null hypothesis for decision making;
5) Output the result if a change is detected; otherwise, go to Step 2 to continue.

V. EXPERIMENTAL VALIDATION

In this section, we will first investigate the impacts of different similarity metrics on EEG change detection by the proposed method, and then compare our method with state-of-the-art methods.

Two different data sets are used in this section, described as below.

- The first EEG data set is taken from the public Bern-Barcelona EEG data set [47]. The original data are recorded with a sampling rate of 1,024 Hz. They randomly select 3,750 pairs of simultaneously recorded signals from the pool of all signals measured at focal and non-focal EEG channels respectively, which means that there are 3750 pieces of data collected from focal area and 3750 pieces of data collected from non-focal area. Then divide the recordings into fragments of length 10000, about 9.8 seconds. The testing dataset used in this experiment contains 350 signals which are formed by concatenating a non-focal signal record and a focal signal record, such that at least one change can be contained in each data stream. Therefore, the length of each piece of signal is about 19.5 seconds, among which 4 seconds data is used for training and initialization.
- The second EEG data set is collected from the Department of Neurology, Second Hospital of Shandong University. The system is shown in Fig. 4. It is used to monitor epilepsy patient in night sleeping. The EEG record data and corresponding surveillance video frames, are collected simultaneously from the monitored individual. The raw EEG signals are recorded with a sampling rate of 1,024 Hz, and downsampled to 512 Hz prior to further analysis. Then it will be sent to PC for analysis. The testing data in this section is taken from channel C4 lasting 30 minutes where the data for 1 minute at the beginning are used as training and initialization data. The neurologist first check the collected data including the EEG and the video, and then label the changes in the data. With this, we can get a number of change points to evaluate the performance of proposed method.[1]

We can use three statistical indicators: *Precision*, *Recall* and comprehensive indicator F_{score}, to evaluate the detection performance, which are defined respectively as below.

$$Precision = \frac{N_1}{N_2} \tag{21}$$

$$Recall = \frac{N_1'}{N} \tag{22}$$

$$F_{score} = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{23}$$

where N_1 is the number of detected points that are true changes, N_2 is the number of all detected points, N_1' is the number of changes that are detected, and N is the number of changes in real situations. An example is given in Fig. 5. There are totally five points that are detected as changes, however, only two of them are in the abnormal status. Since the proposed framework starts with the normal EEG status, that is we use the EEG data under normal status as the 'regular' data, the detected change points that are in normal EEG status will be considered as 'wrong'; in comparison, the detected change points that are in abnormal EEG status will be considered as 'right'. As a result, the statistical indicators are obtained as $N_1 = 2, N_2 = 5, N_1' = 1, N = 2$, and the *Precision* is 0.4, the *Recall* is 0.5, and the F_{score} is 0.44, accordingly

[1]The interested reader can contact the corresponding author for acquiring the testing data.

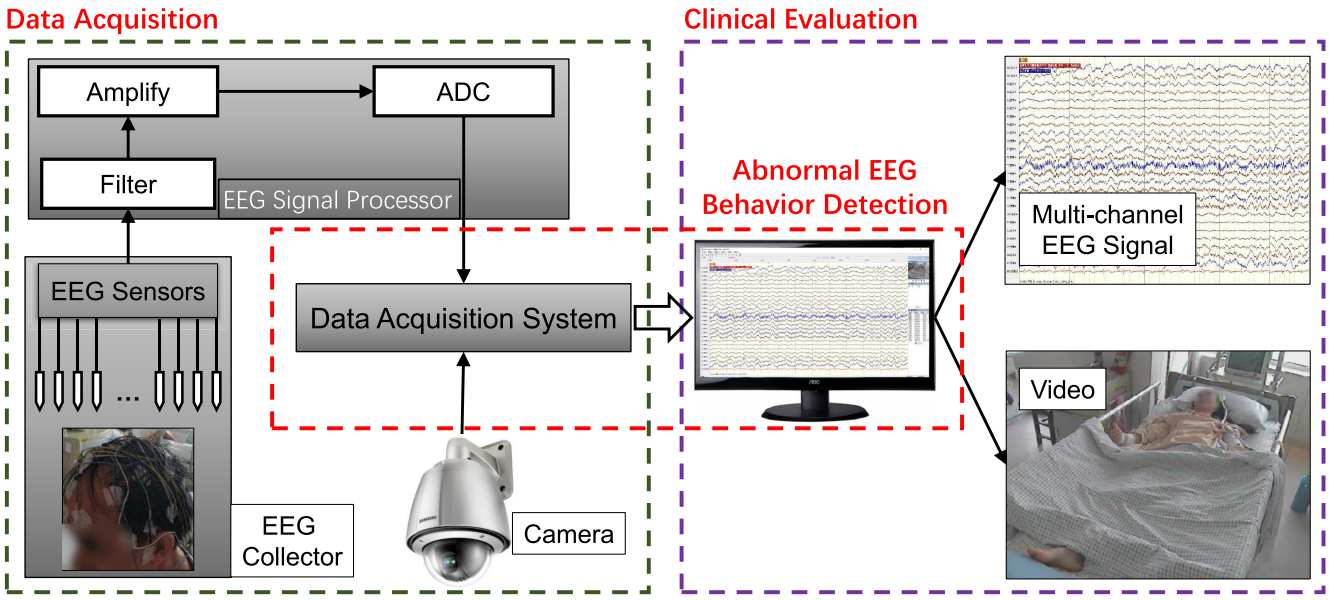


FIGURE 4. The system setup for night sleep monitoring.

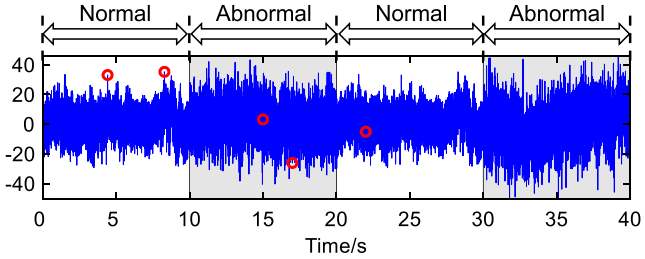


FIGURE 5. The example to explain the statistical indicators where all detected changes are labeled by red circles.

Apparently, a high value of *Precision* means a high detection accuracy, and a high value of *Recall* indicates a high sensitivity of detection, the indicator F_{score} gives a comprehensive evaluation of a detection method. Here, it is worth mentioning that, in the following experiment, once a change is detected, the proposed method continues until the whole testing data is checked. The detected change points are used to calculate the performance indicators where the detected point in abnormal status is as true alarm and the detected point in normal status is as false alarms as the detection process begins with a normal status. The points that are detected as normal is used to update the parameters in Gaussian model in Eq.4.

A. DETECTION PERFORMANCE BY INVESTIGATED METRICS

In this experiment, six metrics introduced in Section III-B are investigated.

- *Bern-Barcelona EEG data set*: First, we show three examples of detected result in Fig. 6. It can be seen that, ED and SKLD seems to be ineffective for the testing data as they did not detect changes accurately. PCCD and KD can detect these changes but with a number of false alarms. HD detected the changes of all testing data but having false alarms for the detection of third testing data. BD performs the best as it can detect the changes accurately for all three testing data. Quantitative comparison is then given in Table 2 for all testing data in the data set. More clearly, it can be seen that, BD and HD perform better than the others. Based on thee results, the investigated metrics can be ranked as BD > HD > PCCD > KD > SKLD > ED.
- *Our testing data set*: Second, for the testing data collected by our system setup, we show the detailed detection results in Fig. 7. For metrics of ED, PCCD, SKLD, KD, many false alarms are generated. Similar to the observation in results on Bern-Barcelona EEG data set, the metrics of HD and BD performs the best, although there are some false alarms in HD. A quantitative comparison of all methods is given in the Table 3. Clearly one can see that, BD achieves the best for all terms of *Precision*, *Recall* and F_{score}, i.e., all three indicators achieve 1. The performance of HD is second only to BD. Therefore, all investigated metrics can be ranked as BD > HD > PCCD > KD > SKLD > ED for our testing data set.

B. COMPARISON WITH STATE-OF-THE-ARTS

To investigate the priority of our method, state-of-the-art methods including Mean, Root Mean Square (RMS), Kurtosis, Skewness, STFT, Empirical Mode Decomposition (EMD), Continuous Wavelet Transform (CWT), Discrete Wavelet Transformation (DWT), method of combining power spectrum and center frequency (CF), method of combining power spectrum and frequency variance (FV) were used to compare. We use them to calculate the anomaly score for the

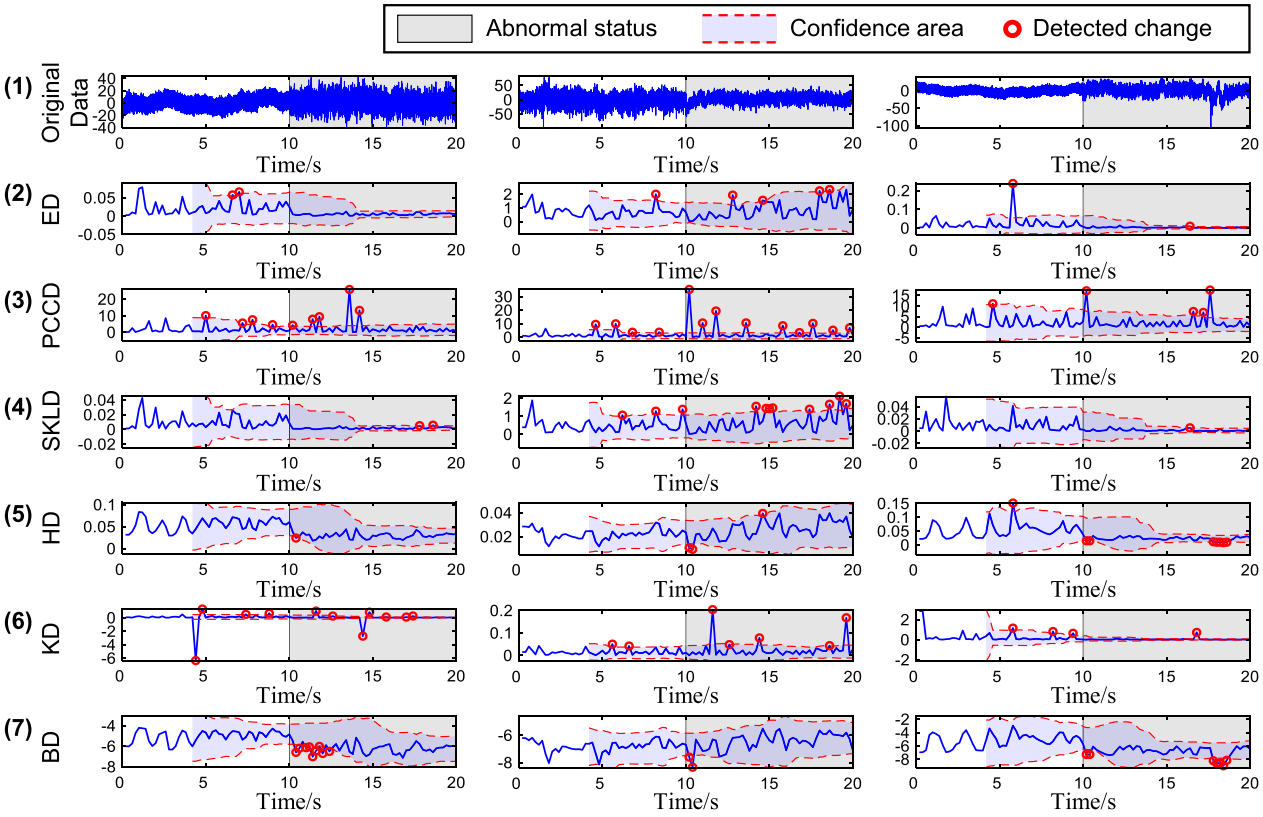


FIGURE 6. Three examples of detected result for testing data. (1): original EEG data, (2)-(7): detected results by ED, PCCD, SKLD, HD, KD and BD, respectively. For each of them, from the left to right are the results for the first testing data, the second testing data and the third testing data.

TABLE 2. Comprehensive comparison in Bern-Barcelona data set. The bold font is used to emerge the best result.

	ED	PCCD	SKLD	HD	KD	BD
$Precision$	0.2857	0.6351	0.3500	**1.0000**	0.5593	0.8868
$Recall$	0.2400	**0.9400**	0.2800	0.7800	0.6600	**0.9400**
F_{score}	0.2609	0.7581	0.3111	0.8764	0.6055	**0.9126**

TABLE 3. Comprehensive comparison in our testing data set. The bold font is used to emerge the best result.

	ED	PCCD	SKLD	HD	KD	BD
$Precision$	0.4786	0.5909	0.4868	0.8889	0.6031	**1.0000**
$Recall$	0.7500	**1.0000**	0.7500	0.8750	0.8750	**1.0000**
F_{score}	0.5843	0.7429	0.5904	0.8819	0.7140	**1.0000**

testing EEG signals and detect the change also by testing the null hypothesis described in Section III for a fair comparison. Moreover, since the BD metric has been demonstrated the priority than others in the above experiment, we directly use it in this experiment.

- *Bern-Barcelona EEG data set*: First, we also show the detection results for three examples used in Fig.8. It can be seen that our method can detect all the changes in three testing data. While, the method using Mean did not detect all the change points, and meanwhile it has a large detection delay even for a successful detection. RMS did not detect the change for the first testing data. Kurtosis and Skewness can detect all changes accurately in the three testing data but there is a large detection delay in the result of Kurtosis. The methods of STFT, EMD, CWT and DWT has a very similar detection performance, that is, they can keep stable in normal EEG status but changes greatly after a change occurs. But the main disadvantage of them is the false detection, i.e., they produce more false alarms than our method. CF and

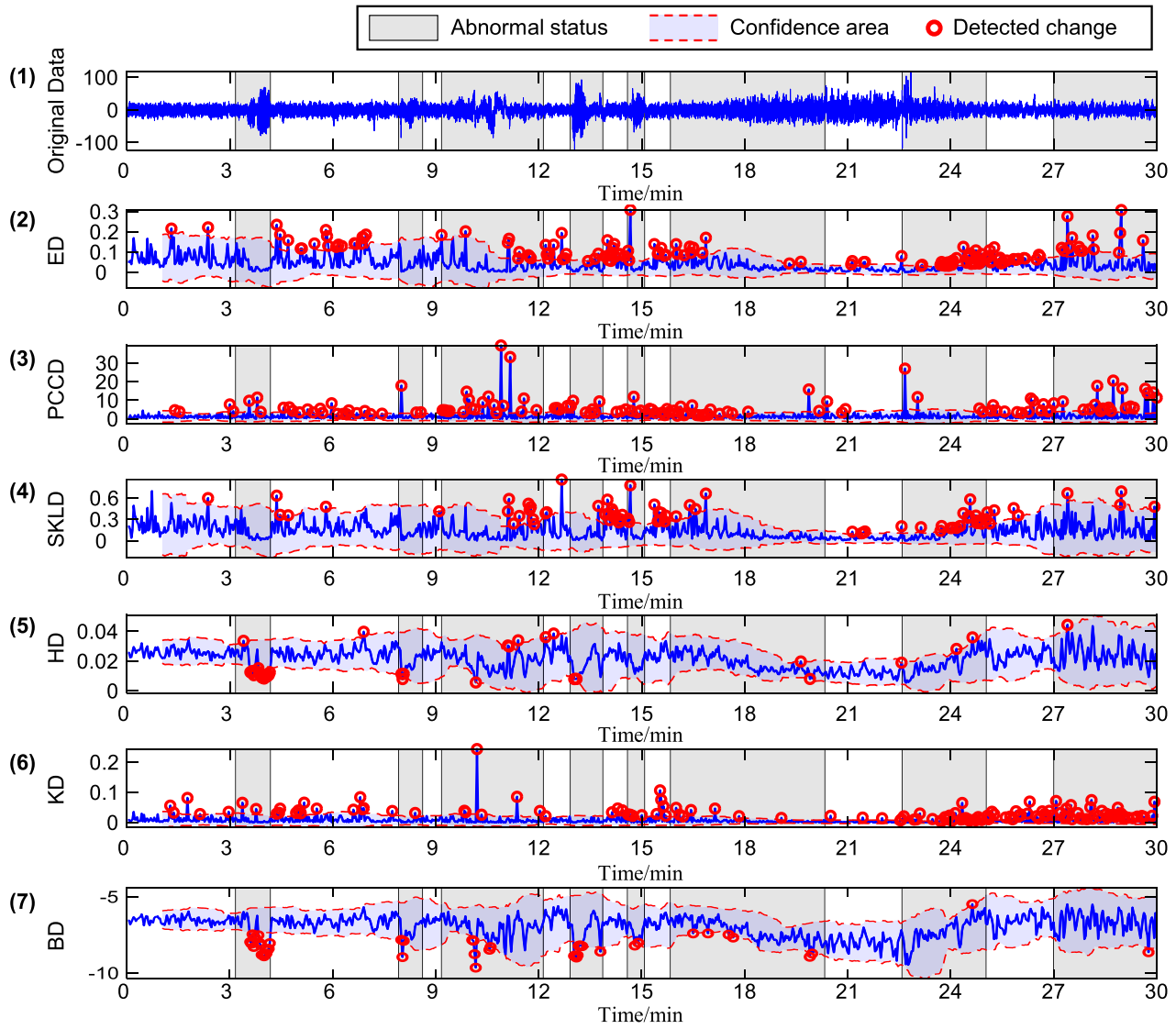


FIGURE 7. Detected results of our testing data. (1): original EEG data, (2)-(7): detected results by ED, PCCD, SKLD, HD, KD and BD, respectively.

FV seems ineffective for all testing data. Comprehensive comparison for all testing data is provided in Table 4 where it can be clearly seen that the proposed method achieves the best performance F_{score} although it has a little lower *Precision* than Skewness and CWT and a little lower *Recall* than EMD.

- *Our testing data set*: Second, the detailed detection results for the testing data collected by our system setup are shown in Fig. 9. It is obvious that our method accurately detects all change points. RMS, STFT, EMD, DWT and FV also detect all change points but meanwhile having a number of false alarms. Comprehensive comparison of all methods is shown in Table 5. It can be seen that the proposed method achieves the best performance among all compared method, revealing the effectiveness of great potential of our method in real applications.

C. RESULT SUMMARY

Combining the results in two above experiments, we found that,

1) The metrics of BD and HD outperform other alternatives, i.e., these two metrics are more appropriate for EEG change detection. However, it is hard to determine which of them is better because their calculations are based on the same transformations of the Bhattacharyya coefficient $BC(P, Q)$, i.e.,

$$\begin{aligned} s^{(HD)} &= 1 - BC(P, Q), \\ s^{(BD)} &= -\ln(BC(P, Q)). \end{aligned} \tag{24}$$

In this regard, HD and BD can be thought of as an approximately equivalent measurement of two statistical samples. The difference between them is the sensitivity to noise as discussed in [48]. Thus, we recommend the

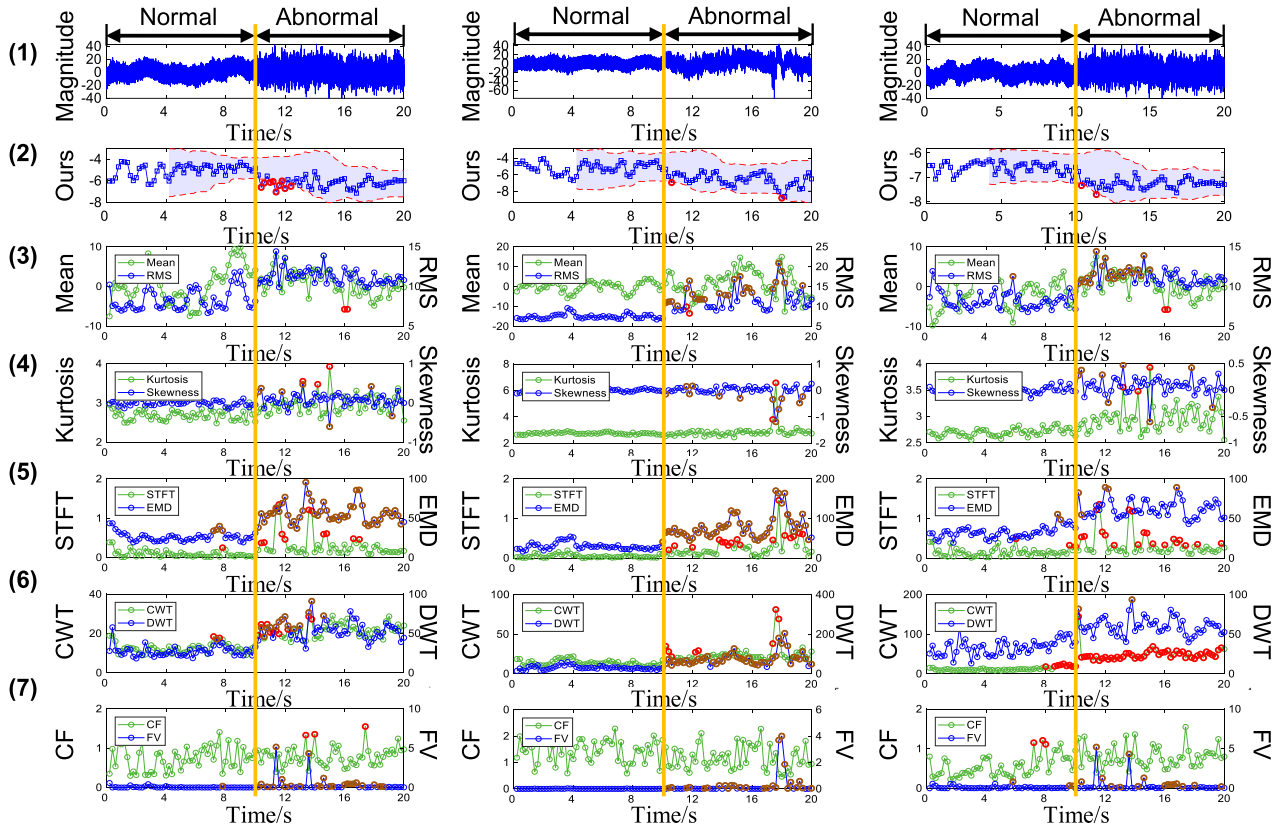


FIGURE 8. Detection results for three examples of testing data where all detected changes are labeled by red or brown circles. (1)-(7) are the original EEG data, the detection results by our proposed method, Mean and RMS, Kurtosis and Skewness, STFT and EMD, CWT and DWT, CF and FV. For each of them, from the left to right are the results for the first testing data, the second testing data and the third testing data.

TABLE 4. Comprehensive comparison in Bern-Barcelona data set. The bold font is used to emerge the best result.

	our proposed method	Mean	RMS	Kurtosis	Skewness	STFT	EMD	CWT	DWT	CF	FV
$Precision$	0.8868	0.6534	0.6164	0.7647	**0.9318**	0.6528	0.6806	**0.9318**	0.7600	0.4762	0.6234
$Recall$	0.9400	0.3200	0.8793	0.7655	0.8514	0.9400	**0.9740**	0.8100	0.9200	0.2000	0.9600
F_{score}	**0.9126**	0.4296	0.7247	0.7651	0.8898	0.7705	0.8013	0.8666	0.8324	0.2817	0.7559

TABLE 5. Comprehensive comparison in our testing data set. The bold font is used to emerge the best result.

	our proposed method	Mean	RMS	Kurtosis	Skewness	STFT	EMD	CWT	DWT	CF	FV
$Precision$	**1.0000**	0.7727	0.8624	0.6512	0.8929	0.7200	0.8729	0.8929	0.8991	0.5815	0.6614
$Recall$	**1.0000**	0.7500	**1.0000**	0.8750	0.7500	**1.0000**	**1.0000**	0.7500	**1.0000**	0.6250	**1.0000**
F_{score}	**1.0000**	0.7612	0.9261	0.7467	0.8152	0.8372	0.9321	0.8152	0.9469	0.5668	0.7962

use of BD or HD in the real implementation of our method in real applications.

2) The main challenge of EEG signal analysis is mostly due to the high noise. The proposed method shows a higher ability to absorb data fluctuations caused by noise, as such it can achieve the best detection performance than state-of-the-arts. We should note that the time-frequency methods including STFT, EMD, CWT and DWT can also show a high robustness to noise (see their detection results in Fig.8), however they have a number of false alarms (for example, see the detection results for the first and third testing data by EMD, CWT). In this regard, these time-frequency methods can improve their detection results using appropriate post-processing technique to remove these false alarms. Meanwhile, comparison results with CF and FV demonstrate the superiority of the proposed method to them.

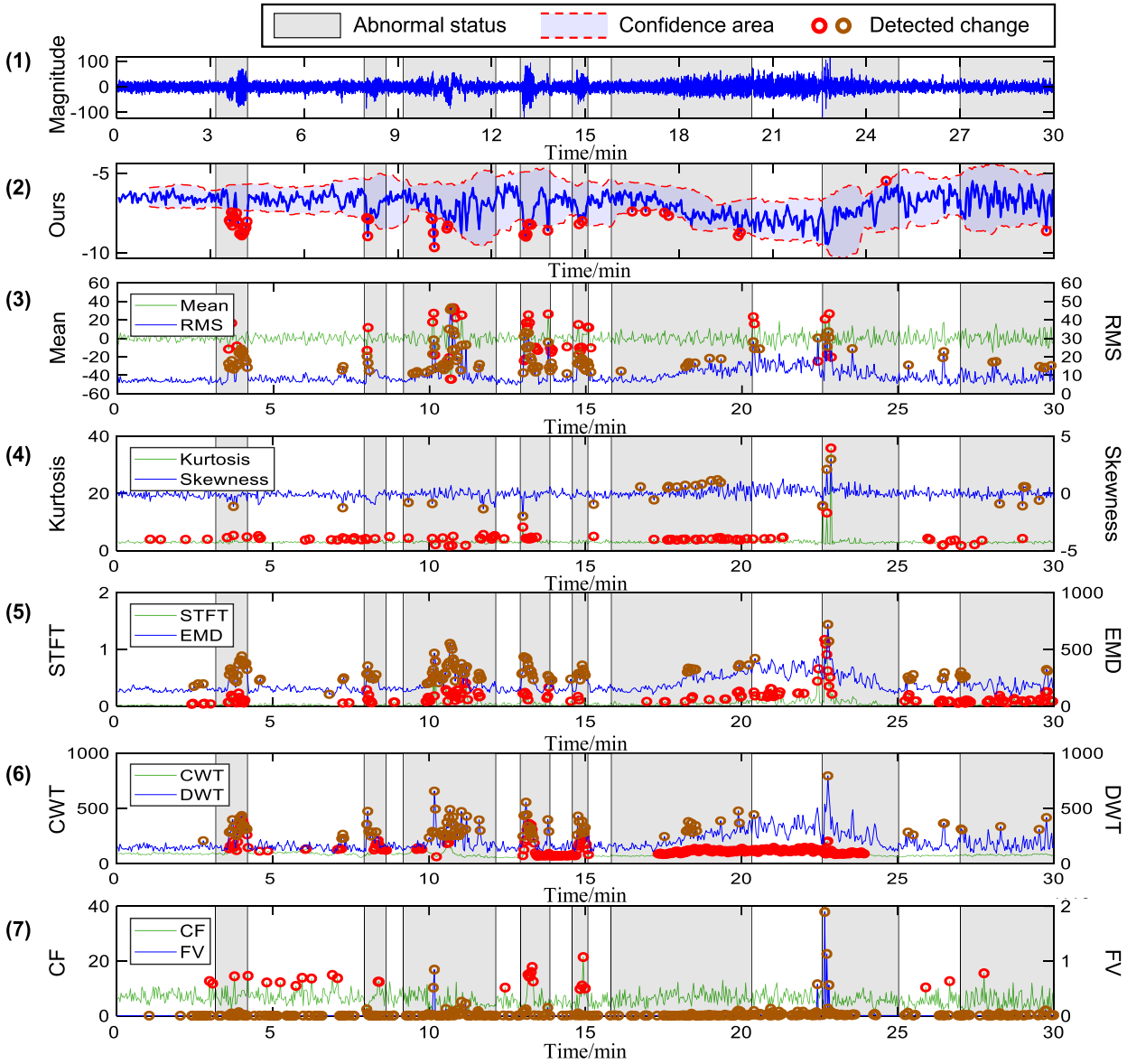


FIGURE 9. Detection results of our testing data. (1)-(7) are the original EEG data, the detection results by our proposed method, Mean and RMS, Kurtosis and Skewness, STFT and EMD, CWT and DWT, CF and FV.

The framework we propose is ideal for real-time monitoring of changes in EEG signals. This is mainly because this method compares the spectral similarity of adjacent segment signals, and can effectively detect the changes of EEG signals at various frequencies. The other compared methods only show the time domain, frequency domain, and time-frequency domain characteristics of the EEG signal, and do not separately calculate the feature changes. Correspondingly, this framework uses the similarity measure to calculate the degree of change, so it is more suitable for change detection for EEG signals. Our proposed framework has achieved good results using BD as the similarity metric, which can prove that, The effect of BD is the best, which shows that BD is the most suitable metric for measuring the similarity of EEG signals in the framework of our method.

It should be noticed that the most commonly used ED did not achieve good results, which means that it is not feasible to simply measure the similarity from the perspective of amplitude. In addition, HD is the best indicator of performance other than BD. The main reason is that both BD and HD are obtained by certain transformations of the Bhattacharyya coefficient $BC(P, Q)$ as Eq. 24. In this regard, HD and BD are thought of as an approximately equivalent measurement of two statistical samples. The difference between them is the sensitivity to noise as discussed in [48]. And the results prove that the sensitivity of BD is more suitable for EEG signal.

In addition, the effect of PCCD is also good, but the other three indicators are very poor.

Therefore, our proposed method based on BD can monitor the state of EEG signals in real time, and can detect changes in an accurate and timely manner.

VI. CONCLUSION

This paper has presented a new method to detect the change in EEG signals based on structural time-series analysis. The method uses a non-overlapping sliding window strategy quantifying the temporal similarity of considered EEG data and detects the change by testing a null hypothesis. The main challenge of this idea is to design an appropriate distance metric that is compatible with our considered data/problem. To achieve this end, we collect a variety of metrics from other areas and investigate them for the problem of EEG change detection. Based on comprehensive experiments conducted on two data sets, it is concluded that, 1) the metric of BD and HD are more appropriate for EEG signal processing; 2) the proposed method outperforms state-of-the-art methods. In future work, we will explore post-processing method to remove false alarms.

REFERENCES

[1] P. Li, C. Karmakar, C. Yan, M. Palaniswami, and C. Liu, "Classification of 5-s epileptic eeg recordings using distribution entropy and sample entropy," *Frontiers Physiol.*, vol. 7, no. 66, p. 136, Apr. 2016.

[2] L. M. Hively, V. A. Protopopescu, and P. C. Gailey, "Timely detection of dynamical change in scalp EEG signals," *Chaos*, vol. 10, no. 4, pp. 864–875, 2000.

[3] M. F. M. Saaid, W. A. B. W. Abas, H. Aroff, N. Mokhtar, R. Ramli, and Z. Ibrahim, "Change point detection of eeg signals based on particle swarm optimization," in *Proc. 5th Kuala Lumpur Int. Conf. Biomed. Eng.* Springer, 2011, pp. 484–487.

[4] M. Sun, F. Wang, T. Min, T. Zang, and Y. Wang, "Prediction for high risk clinical symptoms of epilepsy based on deep learning algorithm," *IEEE Access*, vol. 6, pp. 77596–77605, 2018.

[5] M. Diykh, Y. Li, and P. Wen, "EEG sleep stages classification based on time domain features and structural graph similarity," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 11, pp. 1159–1168, Nov. 2016.

[6] M. S. J. Solaija, S. Saleem, K. Khurshid, S. A. Hassan, and A. M. K. Kamboh, "Dynamic mode decomposition based epileptic seizure detection from scalp EEG," *IEEE Access*, vol. 6, pp. 38683–38692, 2018.

[7] S. Supriya, S. Siuly, H. Wang, J. Cao, and Y. Zhang, "Weighted visibility graph with complex network features in the detection of epilepsy," *IEEE Access*, vol. 4, pp. 6554–6566, 2016.

[8] A. Sharmila and P. Geethanjali, "DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers," *IEEE Access*, vol. 4, pp. 7716–7727, 2016.

[9] G. Wang, Z. Deng, and K.-S. Choi, "Detection of epilepsy with electroencephalogram using rule-based classifiers," *Neurocomputing*, vol. 228, pp. 283–290, Mar. 2017.

[10] Y. Shin, S. Lee, M. Ahn, H. Cho, S. C. Jun, and H.-N. Lee, "Noise robustness analysis of sparse representation based classification method for non-stationary EEG signal classification," *Biomed. Signal Process. Control*, vol. 21, pp. 8–18, Aug. 2015.

[11] Z. Gao, G. Lu, P. Yan, C. Lyu, X. Li, W. Shang, Z. Xie, and W. Zhang, "Automatic change detection for real-time monitoring of EEG signals," *Frontiers Physiol.*, vol. 9, p. 325, Apr. 2018.

[12] J. G. Bogaarts, D. M. W. Hilkman, E. D. Gommer, V. H. J. M. van Kranen-Mastenbroek, and J. P. H. Reulen, "Improved epileptic seizure detection combining dynamic feature normalization with EEG novelty detection," *Med. Biol. Eng. Comput.*, vol. 54, no. 12, pp. 1883–1892, 2016.

[13] A. Gardner, A. Krieger, G. Vachtsevanos, and B. Litt, "One-class novelty detection for seizure analysis from intracranial EEG," *J. Mach. Learn. Res.*, vol. 7, pp. 1025–1044, Jun. 2006.

[14] E. R. Faria, I. J. C. R. Gonçalves, A. C. P. L. F. de Carvalho, and J. Gama, "Novelty detection in data streams," *Artif. Intell. Rev.*, vol. 45, no. 2, pp. 235–269, 2016.

[15] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Process.*, vol. 99, pp. 215–249, Jun. 2014.

[16] I. Bukovsky, M. Cejnek, J. Vrba, and N. Homma, "Study of learning entropy for onset detection of epileptic seizures in EEG time series," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 3302–3305.

[17] Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, and A. Cichocki, "Sparse Bayesian classification of EEG for brain–computer interface," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 11, pp. 2256–2267, Nov. 2016.

[18] D. K. Ravish, S. S. Devi, S. G. Krishnamoorthy, and M. R. Karthikeyan, "Detection of epileptic seizure in EEG recordings by spectral method and statistical analysis," *J. Appl. Sci.*, vol. 13, no. 2, pp. 207–219, 2013.

[19] Z. Gu, G. Yan, J. Zhang, Y. Li, and Z. L. Yu, "Automatic epilepsy detection based on wavelets constructed from data," *IEEE Access*, vol. 6, pp. 53133–53140, 2018.

[20] A. S. M. Murugavel and S. Ramakrishnan, "Hierarchical multi-class SVM with ELM kernel for epileptic EEG signal classification," *Med. Biol. Eng. Comput.*, vol. 54, no. 1, pp. 149–161, Jan. 2016.

[21] J. Yang, H. Singh, E. L. Hines, F. Schlaghecken, D. D. Iliescu, M. S. Leeson, and N. G. Stocks, "Channel selection and classification of electroencephalogram signals: An artificial neural network and genetic algorithm-based approach," *Artif. Intell. Med.*, vol. 55, no. 2, pp. 117–126, 2012.

[22] B. Boashash and S. Ouelha, "Automatic signal abnormality detection using time-frequency features and machine learning: A newborn EEG seizure case study," *Knowl. Based Syst.*, vol. 106, pp. 38–50, Aug. 2016.

[23] H. R. Al Ghayab, Y. Li, S. Siuly, and S. Abdulla, "A feature extraction technique based on tunable Q-factor wavelet transform for brain signal classification," *J. Neurosci. Methods*, vol. 312, pp. 43–52, Jan. 2019.

[24] S. M. Mohammadi, S. Kouchaki, M. Ghavami, and S. Sanei, "Improving time–frequency domain sleep EEG classification via singular spectrum analysis," *J. Neurosci. Methods*, vol. 273, pp. 96–106, Nov. 2016.

[25] J. Xiang, C. Li, H. Li, R. Cao, B. Wang, X. Han, and J. Chen, "The detection of epileptic seizure signals based on fuzzy entropy," *J. Neurosci. Methods*, vol. 243, pp. 18–25, Mar. 2015.

[26] K. Samiee, P. Kovács, and M. Gabbouj, "Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local Gabor binary patterns feature extraction," *Knowl.-Based Syst.*, vol. 118, pp. 228–240, Feb. 2017.

[27] P. Kellaway, "An orderly approach to visual analysis: Characteristics of the normal EEG of adults and children," in *Current Practise in Clinical Electroencephalography*, D. D. Daly and T. A. Pedley, Eds., 2nd ed. New York, NY, USA: Raven Press, 1990, pp. 139–199.

[28] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *J. Neurosci. Methods*, vol. 123, no. 1, pp. 69–87, Feb. 2003.

[29] A. S. M. Murugavel, S. Ramakrishnan, U. Maheswari, and B. S. Sabetha, "Combined seizure index with adaptive multi-class SVM for epileptic EEG classification," in *Proc. Int. Conf. Emerg. Trends VLSI, Embedded Syst., Nano Electron. Telecommun. Syst. (ICEVENT)*, Jan. 2013, pp. 1–5.

[30] G. Chawla, A. Gupta, and S. Sengar, "A wavelet based detection and analysis of gamma rhythms in EEG signals," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Dec. 2004, pp. 22–26.

[31] Y. Jiang, Z. Deng, F.-L. Chung, G. Wang, P. Qian, K.-S. Choi, and S. Wang, "Recognition of epileptic EEG signals using a novel multiview TSK fuzzy system," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 1, pp. 3–20, Feb. 2017.

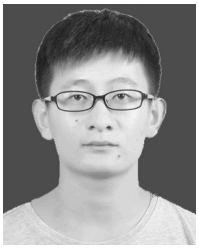
[32] Y. Zhang, S. Yang, Y. Liu, Y. Zhang, B. Han, and F. Zhou, "Integration of 24 feature types to accurately detect and predict seizures using scalp EEG signals," *Sensors*, vol. 18, no. 5, p. 1372, 2018.

[33] A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont, "Predicting epileptic seizures in scalp EEG based on a variational Bayesian Gaussian mixture model of zero-crossing intervals," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 5, pp. 1401–1413, May 2013.

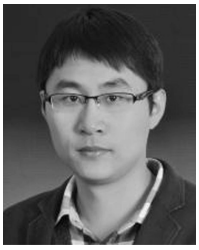
[34] G. Safont, A. Salazar, L. Vergara, E. Gómez, and V. Villanueva, "Multichannel dynamic modeling of non-Gaussian mixtures," *Pattern Recognit.*, vol. 93, pp. 312–323, Sep. 2019.

[35] Z. S. Xu and J. Chen, "An overview of distance and similarity measures of intuitionistic fuzzy sets," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 16, no. 4, pp. 529–555, 2008.

[36] K. Alotaibi, "Non-metric multi-dimensional scaling for distance-based privacy-preserving data mining," Ph.D. dissertation, School Comput. Sci., Univ. East Anglia, Norwich, U.K., 2014.
[37] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "Euclidean distance matrices: Essential theory, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 12–30, Nov. 2015.
[38] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise Reduction Speech Processing*. Berlin, Germany: Springer, 2009, pp. 1–4.
[39] S. Kullback, *Information Theory and Statistics*. New York, NY, USA: Dover, 1997.
[40] S. Tabibian, A. Akbari, and B. Nasersharif, "Speech enhancement using a wavelet thresholding method based on symmetric Kullback–Leibler divergence," *Signal Process.*, vol. 106, pp. 184–197, Jan. 2015.
[41] E. Hellinger, "Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen," *J. für die reine und angewandte Mathematik*, vol. 136, pp. 210–271, Jul. 1909.
[42] A. N. Kolmogorov, "On the approximation of distributions of sums of independent summands by infinitely divisible distributions," in *Contributions to Statistics*. Amsterdam, The Netherlands: Elsevier, 1965, pp. 159–174.
[43] C. A. Fuchs and J. V. D. Graaf, "Cryptographic distinguishability measures for quantum-mechanical states," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1216–1227, May 1999.
[44] H. Laurent and C. Doncarli, "Stationarity index for abrupt changes detection in the time-frequency plane," *IEEE Signal Process. Lett.*, vol. 5, no. 2, pp. 43–45, Feb. 1998.
[45] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, no. 1, pp. 99–109, 1943.
[46] B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi, "A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data," *Knowledge-Based Syst.*, vol. 82, pp. 163–177, Jul. 2015.
[47] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 86, no. 4, 2012, Art. no. 046206.
[48] A. Bhattacharya, P. Kar, and M. Pal, "On low distortion embeddings of statistical distance measures into low dimensional spaces," in *Database and Expert Systems Applications* (Lecture Notes in Computer Science), vol. 5690, no. 3. Berlin, Germany: Springer, 2009, pp. 164–172.



GUANGYUAN CHEN received the bachelor's degree from Shandong University, in 2018, where he is currently pursuing the master's degree. His research interests include signal processing, EEG condition monitoring, and machine health management.



GUOLIANG LU received the bachelor's and master's degrees from Shandong University, Jinan, China, in 2006 and 2009, respectively, and the Ph.D. degree from the Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Japan, in March 2013.

He is currently an Associate Professor with Shandong University. His research interests include signal processing, pattern recognition, and their applications in industrial and biomedical scenarios. He is also a Committee Member of Annual IEEE PHM Conferences, an Associate Editor of the *Journal of Prognostics and Health Management*.

WEI SHANG received the master's degree from Shandong Medical University, Jinan, China, in 1997, and the M.D. degree from Shandong University, Jinan, in 2005.

He is currently a Professor with the Second Hospital of Shandong University. His research interests include epilepsy and sleep disorders. He is also a Committee Member of the American Epilepsy Society, a Committee Member of the Chinese Physicians Association Neurology Branch epilepsy Professional Committee.

ZHAOHONG XIE received the master's and Ph.D. degrees from Shandong University, Jinan, China, in 2004 and 2012, respectively. He is currently an Associate Chief Physician with the Second Hospital of Shandong University. His research interests include electroencephalogram, polysomnography, and their applications in neurology, and sleep medicine.